Essays on transfer pricing, taxes and corporate sustainability performance

Dissertation

submitted to the

Faculty of Business, Economics and Informatics of the University of Zurich

to obtain the degree of

Doktor der Wirtschaftswissenschaften, Dr. oec.

(corresponds to Doctor of Philosophy, PhD)

presented by

Benedikt Bisig

from Einsiedeln, Schwyz

approved in October 2017 at the request of

Prof. Dr. Dieter Pfaff

Prof. Dr. Robert F. Göx

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 25.10.2017
The Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

Acknowledgments

The essays of this cumulative dissertation were written during my time as a research assistant at the Chair of Accounting of Prof. Dr. Dieter Pfaff in Zurich (Department of Business Administration). Therefore, I would like to thank its members for supporting me in achieving the accomplishment of this dissertation.

First of all, to my doctoral thesis supervisor Prof. Dr. Dieter Pfaff: I would like to thank you for welcoming me in a very warm and friendly way at the Chair of Accounting, for your flexibility as my employer, your tireless support and guidance during these four years. Next, I would like to thank Prof. Dr. Robert F. Göx for his preparedness to take on the role of the second reviewer of my thesis, and for his fruitful comments on the essays of this dissertation during our research seminars. I also would like to express a word of thanks to the senior assistant at our chair, Dr. Katrin Hummel: I will always be grateful to you for being a very competent and reliable co-author, a challenging and motivating tutor, and most of all for having a willing ear during my whole time at the Chair of Accounting. Further thanks goes to my former colleagues at the Chair of Accounting: Silvia Chibane, Dr. Patricia Ruffing-Straube, Dr. Ruta Bilkeviciute, Monica Breitwieser, Ute Laun, Dr. Sandro Kürschner, Dr. Peter Ising, Stefan Beeler, Kevin Schuler, and Christian Sidler for their cooperative minds.

Special thanks goes to my family for their support and patience: my father René Bisig, for always pushing me to work on my dissertation while stepping in for me at the office. My mother Hedwig Maria Bisig, for always caring about the wellbeing of my siblings and me. My sister and two brothers, Stefanie, Kilian and Nicolas Bisig, for their encouragement and moral support since our childhood.

Next, I would like to thank my two closest friends, Michael Brandenberger and Mike Agov, for their unique friendship and loyality. Another word of thanks goes to my band colleagues Roger Bucher, Ariel Hasler, Thomas Egli, Simon Kessler, Patrick Aldo Mitidieri, Pascal Reichmuth, Fabian Reichmuth, Alex Schoch, Walter Rickli, Christian Berger, Sandro Schneider and Aurel Hassler for the unforgettable experiences at our concerts.

Last but not least, I take this opportunity to express my deep gratitude to my girlfriend Nora Nussbaumer, for all her love and support.

Table of Contents

1. Introduction

Transfer prices are one of the central topics of controlling and business administration.[1] The history of transfer pricing literature goes back more than a century to the analysis of Schmalenbach in 1903.[2] Although transfer prices have multiple functions and serve a variety of different purposes, traditional research mainly focused on internal coordination (e.g. Hirshleifer, 1956; Coenenberg, 1973; Winter, 1986, Wagenhofer, 1995; Göx, 1999). In the standard setting of analytical literature[3] the transfer price can be seen "as a device for *coordinating the plans and actions of individual decisions makers in decentralized organizations" (Göx & Schiller, 2007, p.673).* While a higher transfer price increases the selling divisions profit, it decreases the profit of the buying unit. At the same time, the amount of units internally traded (and therefore sold on the final market) are affected by the transfer price in a decentralized organization, thereby also determining overall firm profit. If the two divisions involved are further located in different countries under different tax regimes, tax considerations play an additional important role. Shifting profits from the business unit operating in the country having higher tax rates to the unit located in the country with lower taxation raises overall firm after-tax profit. Thus, it becomes clear that a single transfer price cannot completely fulfill multiple objectives at once (Hummel, 2010, p.44), since multiple conflicts between the different functions of a transfer price can arise: for example, in case of excess capacity of a business unit, a transfer price of 0 could be reasonable in terms of the coordination function. However, when it comes to profit allocation, this would lead to severe distortions (Ewert & Wagenhofer, 2008, p.579f.). In addition, a transfer price of 0 would not be tax compliant. Consequently, more recent analytical literature on transfer pricing suggests that a single transfer price cannot optimally achieve internal (coordination) and external (tax compliance and tax burden minimization) objectives simultaneously (Baldenius, Melumad, & Reichelstein, 2004; Choe & Hyde, 2007; Hyde & Choe, 2005; Smith, 2002). This finding has also found its way into management accounting textbooks (e.g., Horngren, Datar, & Rajan, 2015, pp. 883-884; Kaplan & Atkinson, 2014, pp. 454, 463-464; Zimmermann, 2013, pp. 186, 191-198). As a solution, researchers suggest the use of multiple transfer prices for different

[1] See Göx and Schiller (2007) for a more detailed literature review on economic transfer pricing.

[2] Since the post-doctoral thesis of Schmalenbach is not available, we refer to an article containing the most important statements published in 1908/1909. For the Anglo literature, the starting point goes back to Hirshleifer (1956).

[3] We refer to the standard setting with two divisions in a decentralized organization: a producing unit, selling (all) its products to the buying unit, which finally sells them on the final product market.

purposes, in particular, one set of books for tax compliance/tax optimization and another set of books for internal management purposes (Baldenius et al., 2004; Choe & Hyde, 2007). However, this proposition does not seem to have found much acceptance in practice. Consultants emphasize that the use of one set of books may be helpful to signal tax authorities that the transfer pricing system is not driven by a tax optimization strategy (Ernst & Young, 2001, 2003). Even analytical researchers acknowledge that multinational enterprises (MNEs), in practice, use one set of books, "*both for simplicity and in order to avoid the possibility that multiple transfer prices become evidence in any disputes with the tax authorities*" (Baldenius et al., 2004, p. 592). Therefore, research still seems to be challenged suggesting practicable solutions. Meanwhile, the ongoing globalization caused an increasing volume of trade to remain inside MNEs and therefore out of the reach of market forces. In the early 1990s, already 60 percent of international trade took place within MNEs (European commission, 2001, p. 23). More recent estimates suggest that this portion has further increased to 80 percent (UNCTAD, 2013, p.135), emphasizing the practical relevance of transfer pricing nowadays. Such development increases the importance of tax considerations when it comes to the pricing of within-firm transactions. MNEs with cross-border internal transactions have realized the huge tax savings potential arising from profit-shifting using inappropriate transfer prices for tangible or intangible services or goods, a strategically chosen financing strategy or hybrid mismatch arrangements. National governments reacted to this behavior by installing regulations for transfer prices, often following the OECD Transfer Pricing Guidelines (first established in 1979). Transfer prices are subject to these national tax regulations determining the applicability of transfer pricing methods in certain situations. The "arm's length principle" claims that the amount charged by a related party to another must be the same as if the parties were not related. Therefore, an arm's-length price for a product or a service should reflect a price on the open market or simply what an unrelated party would charge for it. Consequently, transfer prices need to be compliant with this regulation in order to avoid penalty payments. Over the last decade, the tax strategies of some of the biggest firms of the world (e.g. Amazon, Apple, Google, Starbucks) attracted huge media attention. For example, the guardian titles in an article released in November 2016 *"Google pays €47m in tax in Ireland on €22bn sales revenue"* (Guardian, 2016). The OECD itself reacted with a 15 action plan called Base Erosion Profit Shifting (BEPS). BEPS refers to "*tax avoidance strategies that exploit gaps and mismatches in tax rules to artificially shift profits to low or no-tax locations"* (OECD, 2017). Multiple actions of the BEPS package address transfer pricing, but some new topics like the digital economy are included, too. Simultaneously to the increased

public attention, tax avoidance also attracted more and more interest of researchers in recent years.[4] Several studies investigate determinants, consequences or the magnitude of tax avoidance in different research settings. Following Hanlon and Heitzman (2010), we define tax avoidance broadly as *"the reduction of explicit taxes"* (p.137).[5] A particular interesting research question within the research field of tax avoidance is its relationship with corporate sustainability performance.[6] Recently, academic research on corporate social responsibility (CSR) has dramatically increased (Huang & Watson, 2015).[7] The broad definition of the United Nations, stating that sustainable development "*seeks to meet the needs and aspirations of the present without compromising the ability to meet those of the future*" (United Nations, 1987, part 1, chapter 1, paragraph 49) suggests that sustainable firms consider both short and long-term externalities. Understanding CSR as corporate *"actions that appear to further some social good, beyond the interests of the firm and that which is required by law"* (McWilliams & Siegel, 2001, p.117) manifests a firm's responsibility towards society. According to Stiglitz & Siu (2016, p.1), *"the first responsibility of any corporation is paying your fair share of taxes"*. Therefore, paying taxes and CSR are related topics. However, the relationship between firms' sustainability performance and tax avoidance behavior remains ambiguous, both theoretically and empirically. Davis et al. (2016) provide large-sample evidence for a positive relationship between corporate social responsibility (CSR) and corporate tax avoidance, thereby supporting traditional economic theories. On the other hand, Lanis and Richardson (2015) and Hoi et al. (2013) provide evidence for a negative relationship which supports the applicability of political theories and the corporate culture perspective.

Against this background, the three essays of this cumulative dissertation address different research questions within the given background. In the first article "Funktionen und Methoden der Verrechnungspreisgestaltung" we look at the relevance of different functions (internal and external) and methods within transfer pricing of Swiss MNEs. The aim of the paper is to provide a descriptive overview of actual transfer pricing systems in practice, contrasting them with analytical literature. Survey results of 53 Swiss MNEs indicate several interesting findings: Overall, firms seem to be very satisfied with their transfer pricing system, although

[4] See Hanlon and Heitzman (2010) for a more detailed literature review of tax research and tax avoidance in accounting.

[5] Therefore, we do not to distinguish between tax-compliant and non-compliant tax avoidance.

[6] See Huang and Watson (2015) for a more detailed literature review of corporate social responsibility research in accounting.

[7] Carroll (2016) provides an overview of the frameworks that are typically discussed in the context of CSR and concludes that the differences between the concepts of CSR and sustainability are negligible.

the large majority relies on one set of books only. With respect to the different functions of transfer prices, tax compliance and tax optimization are the most predominant ones. Regarding internal functions solely, divisional profit determination is the most important one, while the traditional coordination function seems to be rather unimportant. Conflicts between different functions of transfer prices can be observed in practice; the most prevalent conflict exists between internal profit determination and tax optimization. With regard to transfer pricing methods, cost-plus methods are still the predominant ones. The most interesting finding is that firms seem to integrate the transfer pricing system into their management control system, for example by referring to tax-compliant transfer prices when it comes to budgeting or the performance-evaluation of business unit managers. According to analytical literature, this should cause severe problems when using a single set of books.

Following this finding, the second essay "Can the Integration of a Tax Compliant Transfer Pricing System into the Management Control System Be Successful? Yes, It Can!" investigates the perceived success of firms integrating a single tax-compliant transfer price into the management control system. As already pointed out, traditional analytical transfer pricing literature suggests that there exists a conflict between tax compliance (or tax burden optimization) and internal coordination in decentralized organizations and recommends the use of two sets of books (Baldenius et al., 2004; Choe & Hyde, 2007; Hyde & Choe, 2005). We use survey data from 38 tax-compliant Swiss MNEs with cross-border internal transactions and one set of books. Applying partial least squares (PLS) analysis, results suggest that the integration of a tax compliant transfer pricing system into the management control system is perceived to be successful, in particular when the transfer pricing system is transparent and can be revised in the case of fundamental management control problems. This finding contradicts analytical literature. Corroborating these findings, we also find affirmative testimonies from interview data of three MNEs. Overall, our results are consistent with survey-based findings reporting that firms predominantly use the same transfer price for both tax compliance and management control purposes.

The third article "Sustainability Performance and Tax Avoidance – Disentangling the Effects from Operational and Management Sustainability Performance" investigates the relationship between firms' sustainability performance and tax avoidance behavior. Since this relation remained ambiguous in literature so far, both theoretically and empirically, we follow Trumpp et al. (2015) and differentiate between an operational and a management sustainability performance. We argue that this distinction might be a reasonable explanation for the

confounding effects of prior literature. Using a sample of 4'449 (respectively 4'427) firm-year observations of U.S. MNEs, we find that management sustainability performance is positively related to corporate tax avoidance, while operational sustainability performance and tax avoidance are negatively associated. This finding is robust for the social and the environmental dimension solely, too. Our reasoning of disentangling management and operational sustainability performance might provide an explanation for the inconsistent findings of prior studies on the relationship between sustainability performance and tax avoidance.

References

Baldenius, T., Melumad, N. D., Reichelstein, S. (2004). Integrating Managerial and Tax Objectives in Transfer Pricing. *The Accounting Review, 79, 2004, 3*, 591-615.

Carroll, A.B., (2015). Corporate social responsibility: The centerpiece of competing and complementary frameworks. *Organizational Dynamics, 44,* 87-96.

Choe, C., & Hyde, C. E. (2007). Multinational transfer pricing, tax arbitrage and the arm's length principle. *Economic Record, 83*(263), 398-404.

Coenenberg, A. (1973). Verrechnungspreise zur Steuerung divisionaler Unternehmen, in: Wirtschaftswissenschaftliches Studium, 2, 8/9, 373-382.

Davis, A.K., Guenther, D.A., Krull, L.K., Williams, B.M., (2016). Do socially responsible firms pay more taxes? *The Accounting Review, 91(1),* 47-68.

Ernst & Young. (2001). *Transfer pricing 2001 global survey. Making informed decisions in uncertain times*. online: http://www.ey.com, download: 29.06.2017.

Ernst & Young. (2003). *Transfer pricing 2003 global survey. Practices, perceptions and trends in 22 countries*. online: http://www.ey.com, download: 29.06.2017.

European Commission (2001). Company Taxation in the Internal Market, Commission Staff Working Paper, Brüssel 2001, online: http://ec.europa.eu/taxation_customs/sites/taxation/files/docs/body/company_tax_study_en.pdf, request date: 29.06.2017.

Ewert, R., Wagenhofer, A. (2008). Interne Unternehmensrechnung, 7. Auflage, Berlin, Heidelberg u. a. 2008.

Göx, R. F. (1999). Strategische Transferpreispolitik im Dyopol, Wiesbaden 1999.

Göx R. F., Schiller U. (2007). An economic perspective on transfer pricing. In: Chapman C. S., Hopwood A. G., Shields M. D. (Eds.), Handbook of accounting management research: 674-693.

Guardian, (2016). Google pays €47m in tax in Ireland on €22bn sales revenue. online: https://www.theguardian.com/business/2016/nov/04/google-pays-47m-euros-tax-ireland-22bn-euros-revenue, request date: 29.06.2017.

Hanlon, M., Heitzman, S., (2010). A review of tax research. *Journal of Accounting and Economics, 50,* 127-178.

Hirshleifer, J. (1956). On the economics of transfer pricing. *Journal of Business, 29,* 172–184.

Hoi, C.K., Wu, Q., Zhang, H., (2013). Is corpoate social responsibility (CSR) associated with tax avoidance? Evidence from irresponsible CSR activities. *The Accounting Review,* 88(6), 2025-2059.

Horngren, C. T., Datar, S. M., & Rajan, M. V. (2015). *Cost accounting: A managerial emphasis.* (14th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Huang, X. and Watson, L.(2015). Corporate Social Responsibility Research in Accounting. *Journal of Accounting Literature 34,* 1-16.

Hummel, K. (2010). Gestaltungsparameter und Einflussfaktoren von Verrechnungspreissystemen, Baden-Baden 2010.

Hyde, C. E., & Choe, C. (2005). Keeping two sets of books: The relationship between tax and incentive transfer prices. *Journal of Economics & Management Strategy, 14*(1), 165-186.

Kaplan, R., & Atkinson, A. A. (2014). *Advanced management accounting.* (3rd ed.). Harlow, Essex: Person.

Lanis, R., Richardson, G., (2012). Corporate social responsibility and tax aggressiveness: an empirical analysis. *J. Account. Public Policy, 31,* 86-108.

McWilliams, A., & Siegel, D. (2001). Corporate social responsibility: A theory of the firm perspective. *Academy of Management Review, 26(1),* 117−127.

OECD, (2017). Base erosion and profit shifting. online: http://www.oecd.org/tax/beps/, request date: 29.06.2017.

Schmalenbach, E. (1908/1909). Über Verrechnungspreise. *Zeitschrift für handelswissenschaftliche Forschung, 3,* 165–185.

Smith, M. J. (2002). Ex ante and ex post discretion over arm's length transfer prices. *The Accounting Review, 77*(1), 161-184.

Stiglitz, J. E., Siu, E. (2016). The Apple tax tussle shows the need for a new way of taxing profits. online:

https://www8.gsb.columbia.edu/faculty/jstiglitz/sites/jstiglitz/files/Apple%20Tax%20Tussle_0.pdf, request date: 29.06.2017.

Trumpp, C., Endrikat, J., Zopf, C., Guenther, E., (2015). Definition, conceptualization, and measurement of corporate environmental performance: A critical examination of a multidimensional construct. *Journal of Business Ethics, 126(2),* 185-204.

United Nations (1987). Report of the world commission on environment and development: Our common future. online: http://www.un-documents.net/our-common-future.pdf, request date: 29.06.2017.

United Nations Conference on Trade and Development (2013). World Investment Report 2013, online: http://unctad.org/en/PublicationsLibrary/diae2013d1_en.pdf, request date: 29.06.2017.

Wagenhofer, A. (1995). Verrechnungspreise zur Verhaltenssteuerung in dezentralisierten Unternehmen. In: Kostenrechnungspraxis, 39, 5, 269-275.

Winter, R. (1986). Pretiale Lenkung bei sicheren und unsicheren Erwartungen, Frankfurt am Main 1986.

Zimmermann, J. L. (2013). *Accounting for decision making and control.* (8th ed.). New York, NY: McGraw-Hill/Irwin.

2.1. Funktionen und Methoden der Verrechnungspreisgestaltung

Prof. Dr. Dieter Pfaff ist seit 1994 Ordinarius für Betriebswirtschaftslehre der Universität Zürich und Inhaber des Lehrstuhls für Accounting, insbesondere Unternehmensrechnung und Controlling. 2011 bis 2014 leitete er zudem als Direktor das Institut für Betriebswirtschaftslehre der Universität Zürich. Davor war er an der Johann-Wolfgang-Goethe-Universität Frankfurt am Main tätig. Er ist Mitherausgeber der Zeitschrift «Die Unternehmung», Mitglied des Editorial Board der «Management Accounting Research», Verwaltungsratsmitglied und Präsident des Beirats der Controller Akademie AG, Zürich, sowie Vizepräsident von veb.ch, des grössten Schweizer Verbands in Rechnungslegung und Controlling. Dieter Pfaff ist Autor und Mitherausgeber mehrerer Fachbücher sowie zahlreicher Beiträge in nationalen und internationalen Fachzeitschriften.

Dr. Katrin Hummel ist Oberassistentin und Habilitandin am Lehrstuhl für Accounting, insbesondere Unternehmensrechnung und Controlling, der Universität Zürich. Zuvor war sie als Controllerin bei einem Konzern im Anlagenbau für das Controlling von internationalen Grossprojekten verantwortlich. Von 2005 bis 2008 arbeitete Katrin Hummel als wissenschaftliche Mitarbeiterin am Lehrstuhl Controlling der Universität Stuttgart. Katrin Hummel ist Autorin verschiedener Beiträge zu den Themen Verrechnungspreisgestaltung, Nachhaltigkeitsberichterstattung und Responsible Management Education.

Benedikt Bisig arbeitet als Assistent (Teilzeit) am Lehrstuhl für Accounting, insbesondere Unternehmensrechnung und Controlling, der Universität Zürich. Nebenbei arbeitet er als Commercial Manager in einer Handelsfirma, welche Rohstoffe der Düngemittelproduktion vermarktet. Benedikt Bisigs Forschungsgebiet liegt im Bereich der Verrechnungspreise.

Dieser Artikel ist veröffentlicht unter Pfaff, D./Hummel, K./Bisig, B., Funktionen und Methoden der Verrechnungspreisgestaltung, in: Meyer, C./Pfaff, D. (Hrsg.): Finanz- und Rechnungswesen Jahrbuch 2015, Zürich, S. 389-416, 2015.

2.1.1. Einführung

Zu den wichtigsten und meist diskutierten Instrumenten dezentraler Steuerung von Organisationen gehören die Verrechnungs- oder Transferpreise. Der Grund dafür ist, dass Verrechnungspreise das Ergebnis der am internen Lieferungs- oder Leistungsaustausch beteiligten Profit Center in der Regel massgeblich beeinflussen und dabei systemimmanente Interessenkonflikte bestehen: «Während die liefernde oder leistende Einheit einen hohen Verrechnungspreis präferiert, stellt sich die Lage für den abnehmenden Bereich genau entgegengesetzt dar. Bei steuerüberschreitenden Verrechnungspreisen treten hierzu noch die Partikularinteressen der Fisci sowie der Zollbehörden der betroffenen Länder.» [1]

In dieser Situation gibt es weder wissenschaftlich eindeutig präferierte, noch einfache praktische Lösungen. Dies gilt vor allem deshalb, weil aus der Kooperation der verbundenen Unternehmen in der Regel Synergien entstehen, die über den Verrechnungspreis alloziiert werden (müssen). Synergien lassen sich aber nur willkürlich aufteilen. Unternehmen versuchen, das Problem pragmatisch zu lösen, indem sie einerseits den Verrechnungspreis so wählen, dass sie sich innerhalb der national oder international steuerlich vorgegebenen Verrechnungspreisrichtlinien bewegen, andererseits aber stets das Gesamtoptimum des Unternehmens im Auge behalten. Die Umsetzung dieses Grundsatzes ist schwierig, da das lokale Management in einem permanenten Zielkonflikt zwischen Bereichs- und Gruppeninteressen steht. Dass das Bereichsinteresse dennoch häufig im Vordergrund steht, hat damit zu tun, dass Bereichsergebnisse in der Berichterstattung häufig zu einseitig herausgestellt werden und zudem gruppenkonformes Verhalten zu wenig honoriert wird.[2] «Bereichsegoismen können aber ein Unternehmen – je nach Grössenordnung – Ergebnisse in Millionenhöhe kosten, eben in Höhe der Deckungsbeiträge, die man an Dritte verschenkt (bei Zukauf statt Eigenbezug) oder erst gar nicht realisiert, weil man auf ein angeblich unrentables Geschäft im Verbund verzichtet.»[3]

Vor diesem Hintergrund kommt der Ausgestaltung des Verrechnungspreises eine zentrale Rolle zu. Der oberste Grundsatz von Verrechnungspreissystemen – zumindest aus steuerlichen Erwägungen heraus – ist nach wie vor der Fremdvergleichsgrundsatz: Verrechnungspreise sollen wie am Markt, also wie gegenüber Dritten gebildet werden («dealing at arm's length»):[4] In der klassischen Verbundbeziehung zwischen einer Produktions- und einer Vertriebsgesellschaft z.B. bedeutet der Fremdvergleichsgrundsatz, dass man der Konzernvertriebsgesellschaft dieselben Konditionen einräumt, wie man dies gegenüber einem unabhängigen Dritten getan hätte.[5] Gemäss den OECD-Verrechnungspreisrichtlinien von 1995 galt bezüglich der Anwendung und Umsetzung des Fremdvergleichsgrundsatzes eine strenge Hierarchie der Verrechnungspreismethoden: So war die Preisvergleichsmethode die bevorzugte Methode, und die Kostenaufschlagsmethode sowie die Wiederverkaufspreismethode sollten den gewinnorientierten Methoden vorgezogen werden. Letztere sollten nur in Ausnahmefälle angewandt werden.[6]

1 Pfaff/Stefani (2006), S. 517.

2 Vgl. Peters/Pfaff (2008/2011), S. 132.

3 Peters/Pfaff (2008/2011), S. 132.

4 Vgl. OECD (2011a), S. 33 ff.

5 Vgl. Peters/Pfaff (2008/2011), S. 132.

6 Vgl. OECD (1995/96/97), Tz. 2.49 und 3.49; Wolff (2010), S. 3.

Mit der Neufassung der Verrechnungspreisgrundsätze 2010 wurde die bisherige Hierarchie aufgehoben. Neu soll die Auswahl der Verrechnungspreismethode immer darauf abzielen, «für jeden Fall die am besten geeignete Methode zu finden.»[7] Nur in Situationen, in denen sich Methoden als gleich zuverlässig erweisen, wird weiterhin der Preisvergleichsmethode gegenüber den anderen Standardmethoden und den Standardmethoden gegenüber den Gewinnmethoden Vorrang eingeräumt.[8] Es kann deshalb vermutet werden, dass in den letzten Jahren Gewinnmethoden wie die Gewinnaufteilungsmethode und die Nettomargenmethode an Bedeutung gewonnen haben. Dies kann man insbesondere dann erwarten, wenn die Transaktion durch den Einsatz wichtiger immaterieller Wirtschaftsgüter geprägt ist oder «aufgrund der hohen Verflechtung ihrer Abläufe einzigartig ist.»[9]

Vor diesem Hintergrund ist es das Ziel des vorliegenden Beitrags, die Verrechnungspreisgestaltung Schweizer Konzerne zu beleuchten. In der Schweiz gibt es zwar keine spezifischen Regelungen zur Gestaltung und Dokumentation von Verrechnungspreisen; allerdings wird auf die Verrechnungspreisrichtlinie der OECD als international gültiger Rahmen verwiesen.[10] Zudem müssen sich Schweizer Konzerne wegen ihrer starken Exportorientierung und Auslandsverbundenheit der internationalen Regulierung beugen. Pfaff und Stefani stellen für Schweizer Unternehmen bereits für das Jahr 2005 fest, dass OECD-richtlinienkonforme Verrechnungspreise die Beziehungen zwischen rechtlich selbständigen Einheiten dominieren.[11]

Im Einzelnen stellen sich folgende Fragen, die auch die Gliederung des Beitrags bestimmen:

1. Welche Funktionen dominieren aus praktischer Sicht die Verrechnungspreisgestaltung? Welche Rolle spielen insbesondere die Erfüllung steuerrechtlicher Regelungen und die Steueroptimierung im Vergleich zu den internen Funktionen der Erfolgsermittlung und Steuerung von Entscheidungen (Koordinationsfunktion)?
2. Welche Bedeutung haben die von der OECD empfohlenen Methoden für die in der Schweiz befragten Unternehmen? Sind die geschäftsvorfallbezogenen Gewinnmethoden überhaupt signifikant vertreten oder dominieren nach wie vor die geschäftsvorfallbezogenen Standardmethoden? Gibt es Unterschiede bei der Wahl der Verrechnungspreismethode zwischen den jeweiligen Wertschöpfungsstufen? Wie werden zentrale Dienstleistungen verrechnet?
3. Wie zufrieden sind Unternehmen mit ihrem Verrechnungspreissystem und welche Rolle spielen dabei Konflikte zwischen den Verrechnungspreiszwecken sowie die Integration des Verrechnungspreissystems in das Management Control System des Unternehmens?

Die Beantwortung dieser Fragen erfolgt auf der Grundlage einer empirischen Erhebung der Verrechnungspreisgestaltung in der Unternehmenspraxis; die Befragung wurde 2012 vom Lehrstuhl für Unternehmensrechnung und Controlling der Universität Zürich vorgenommen. Die letzte ausführliche

7 OECD (2011b), S. 65.

8 Vgl. OECD (2011b), S. 66.

9 Vgl. Wolff (2010), S. 4.

10 Vgl. Kreisschreiben Nr. 4 (Besteuerung von Dienstleistungsgesellschaften) der Eidgenössischen Steuerverwaltung ESTV vom 19. März 2004; Kreisschreiben Nr. 9 (Nachweis des geschäftsmässig begründeten Aufwandes bei Ausland-Ausland-Geschäften) der ESTV vom 22. Juni 2005.

11 Vgl. Pfaff/Stefani (2007), S. 212.

Bestandsaufnahme für die Schweiz – vom selben Lehrstuhl durchgeführt – datiert aus dem Jahr 2005 und fällt in eine Zeit, in der es noch die zuvor beschriebene starke Hierarchie der Verrechnungspreismethoden gab.[12] Auch wenn die in 2005 und 2012 gestellten Fragen nicht völlig deckungsgleich sind, lassen sich interessante Aufschlüsse über die Entwicklung der Verrechnungspreissysteme gewinnen.

2.1.2. Untersuchungssample

Die jüngste, vom Lehrstuhl für Unternehmensrechnung und Controlling durchgeführte Untersuchung basiert auf einer Grundgesamtheit von 158 Unternehmen, die zum 15. Juli 2012 an der Schweizer Börse SIX primärkotiert waren und nicht der Banken-, Versicherungs- oder Immobilienbranche angehörten. Die Finanzdienstleistungsbranche wird typischerweise in solchen Untersuchungen ausgeschlossen, da sie sich wegen ihrer geschäftsspezifischen Merkmale fundamental vom Rest des Untersuchungssamples unterscheidet.[13]

Im Zeitraum zwischen Juli und September 2012 wurden die Unternehmen der ausgewählten Grundgesamtheit zunächst telefonisch kontaktiert und die entsprechenden Verrechnungspreisverantwortlichen innerhalb des Unternehmens ermittelt.[14] Sechs Wochen später wurden Unternehmen, die einer Teilnahme zugestimmt aber noch nicht geantwortet hatten, erneut zum Ausfüllen des Online-Fragebogens eingeladen. Dieses Vorgehen ergab eine Gesamtheit von 68 ausgefüllten Fragebögen, was einer Rücklaufquote von 43 Prozent entspricht. 15 Fragebögen mussten allerdings eliminiert werden, da sie in hohem Mass Lücken aufwiesen. Somit liegt der vorliegenden Untersuchung eine Gesamtheit von maximal 53 Fragebögen zugrunde, was einer effektiven (und aussagekräftigen) Rücklaufquote von 34 Prozent entspricht.[15] In der Untersuchung von 2005 lag die Rücklaufquote bei ebenfalls erfreulichen 37 Prozent.[16]

Bei 83 Prozent aller teilnehmenden Unternehmen handelt es sich um die Konzernobergesellschaft, bei 9 Prozent um eine Tochtergesellschaft. Die restlichen 8 Prozent bilden Obergesellschaften einer Zwischenholding, Divisionen einer Rechtseinheit oder Sonstiges. Die durchschnittliche Mitarbeiteranzahl für das Geschäftsjahr 2011 reicht bei einem Durchschnitt von 7'656 Mitarbeiter von 15 bis 123'686. Der durchschnittliche Nettoumsatz (inklusive Innenumsatz) für die Geschäftsperiode 2011 beträgt 3.019 Mrd. CHF, wobei eine grosse Streuung zwischen 18 Mio. CHF und 52 Mrd. CHF beobachtbar ist. Die Gesamtheit der vorliegenden Fragebögen bildet sowohl hinsichtlich des Umsatzes und der Mitarbeiterzahl als auch anderer wichtiger Merkmale in etwa die Struktur der angeschriebenen Grundgesamtheit ab, so dass zumindest kein Hinweis auf eine eingeschränkte Repräsentativität der Umfrage besteht.

12 Vgl. Pfaff/Stefani (2006). Zur Schweizer Verrechnungspreispraxis vgl. auch Weilenmann (1989).

13 Entsprechendes gilt für die Untersuchung von 2005; vgl. Pfaff/Stefani (2007), S. 204.

14 Jede Bezeichnung der Person, des Status oder der Funktion in der vorliegenden Arbeit gilt in gleicher Weise für Männer und Frauen.

15 Die Grundgesamtheit kann je nach Untersuchung in Form der Stichprobengrösse von den genannten 53 abweichen.

16 Vgl. Pfaff/Stefani (2006), S. 518; Pfaff/Stefani (2007), S. 204.

Abbildung 1 veranschaulicht die Branchenzugehörigkeit der Grundgesamtheit der 53 Unternehmen. Am stärksten vertreten ist der Maschinen- und Anlagenbau, dem knapp ein Viertel der Unternehmen angehören. Es folgen «Pharma/Biotechnologie/Medizinaltechnologie» (15 Prozent) und «Handel und Transport» (ebenfalls 15 Prozent, wobei der Löwenanteil mit 11 Prozent auf Handelsunternehmen entfällt). Des Weiteren befinden sich die Branchen «Baugewerbe/Rohstoffe/Chemie», «Konsumgüterindustrie», «Energieversorgung», «Elektronik», «Telekommunikation» und «Sonstiges» unter den Antwortenden. Nicht vertreten sind «IT/Software» und «Automobil». Vergleicht man die Branchenverteilung mit der Untersuchung 2005 ergibt sich ein ähnliches Bild bei leicht angepasster Branchenkategorisierung.[17]

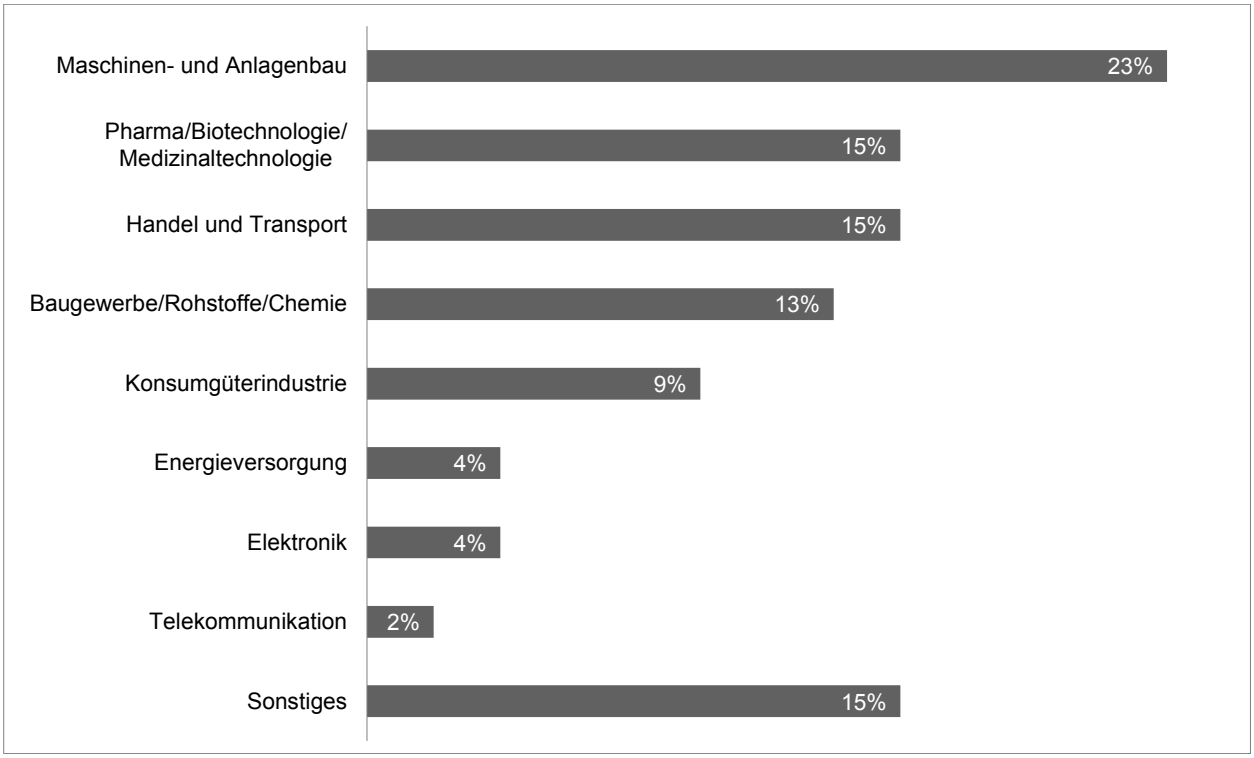


Abbildung.1: Untersuchungssample nach Branchenzugehörigkeit

Zur Illustration des Masses an internationaler Verflechtung wird der Anteil des im Ausland erwirtschafteten Umsatzes (inkl. Innenumsatz) betrachtet. Mit durchschnittlich 66 Prozent liegt dieser sehr hoch und liefert ein deutliches Indiz für die internationale Orientierung von Schweizer Unternehmen. So geben 92 Prozent der Unternehmen an, dass grenzüberschreitender interner Leistungsaustausch vorliegt. Dieser hohe Anteil akzentuiert die Bedeutung von steuerlichen Aspekten bei der Verrechnungspreisgestaltung von Schweizer Unternehmen. Bei 75 Prozent aller Unternehmen findet interner Leistungsaustausch zwischen unterschiedlichen Steuersubjekten innerhalb der Landesgrenzen statt. Nur 35 Prozent aller befragten Unternehmen geben an, dass Leistungsaustausch innerhalb desselben steuerpflichtigen Unternehmens vorkommt.

Abbildung 2 veranschaulicht die Zuständigkeiten bei der Verrechnungspreisgestaltung. Mit einem durchschnittlichen Wert von 4.9 (auf einer Skala von 1 bis 7) ist die zentrale Controllingabteilung bei der

17 Vgl. Pfaff/Stefani (2007), S. 205.

Ausgestaltung der Verrechnungspreise am stärksten eingebunden, gefolgt von der zentralen Steuerabteilung (4.6), der Geschäftsleitung (4.4) und anderen Abteilungen (4.1). In den meisten Fällen werden unterschiedliche Abteilungen in den Entscheidungsprozess eingebunden. Ein Muster an häufig auftretenden Kombinationen lässt sich allerdings nicht erkennen. Bemerkenswert ist also, dass trotz der starken (und zunehmenden) Bedeutung der steuerlichen Compliance (vgl. *Abbildung 4*) der Controllingabteilung – zumindest im Durchschnitt aller Unternehmen – nach wie vor die grösste Rolle zukommt.

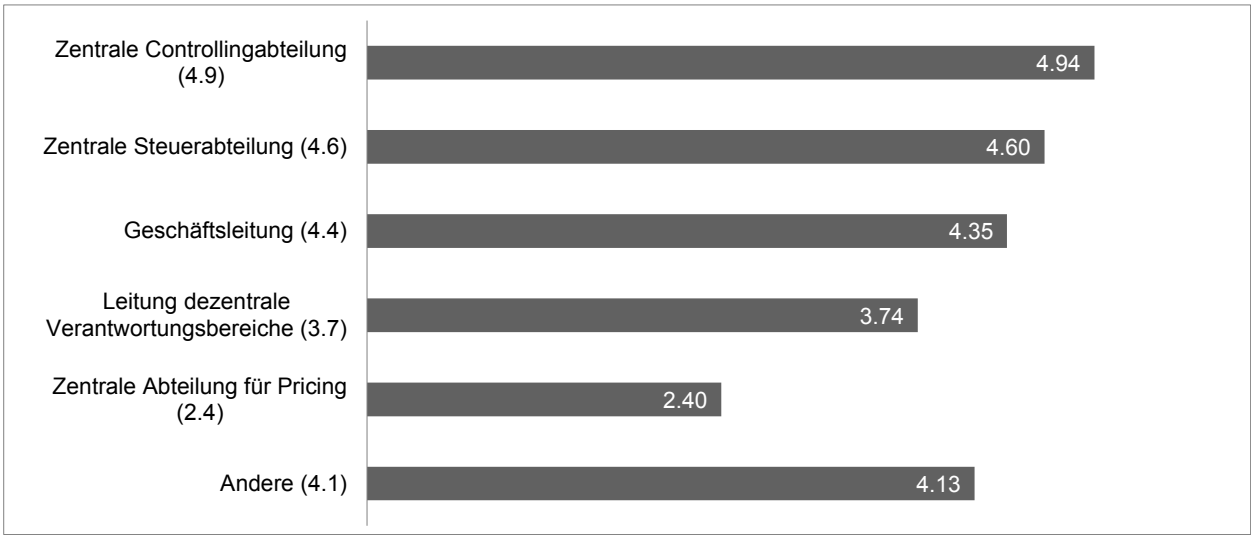


Abbildung 2: Zuständigkeiten bei der Verrechnungspreisgestaltung

2.1.3. Funktionen der Verrechnungspreisgestaltung

Generell unterscheidet man zwischen internen und externen Funktionen von Verrechnungspreisen. Während interne Funktionen vor allem der Erfüllung von unternehmensinternen Zwecken wie beispielsweise der Erfolgsermittlung einzelner Profit Center sowie der dezentralen Steuerung (Koordination) dienen, bilden externe Funktionen die Informationsgrundlage für aussenstehende Stakeholder (z.B. Steuer- und Regulierungsbehörden).[18]

Abbildung 3 veranschaulicht die Bedeutung verschiedener interner Funktionen der Verrechnungspreisgestaltung.[19] Die grösste Bedeutung kommt dabei der Erfolgsermittlung einzelner Profit Center mit einem durchschnittlichen Wert von 4.1 (auf einer Skala von 1 bis 7) zu, dicht gefolgt von der Transparenz unternehmerischer Prozesse und Entscheidungen (4.0) sowie der Vereinfachung der Planung (3.9). Die klassische und in der wissenschaftlichen Literatur oft diskutierte Funktion der Koordination und Beeinflussung der Entscheidungsfindung ist für die Unternehmen des Untersuchungssamples nur von relativ geringer Bedeutung.[20]

18 Vgl. ausführlicher Pfaff/Hummel 2014, S. 593 f.

19 Der ursprünglich ausgehändigte Fragebogen basiert auf einer siebenstufigen Likertskala (1=trifft nicht zu, 7=trifft voll zu). Zur besseren Veranschaulichung wurden die sieben Antwortkategorien zu fünf zusammengefasst: 1=keine Bedeutung, 2+3=geringe Bedeutung, 4=mittlere Bedeutung, 5+6=grosse Bedeutung, 7=sehr grosse Bedeutung.

20 Vgl. bereits Schmalenbach (1909), Hirshleifer (1956) sowie Ronen/McKinney (1970) als die grundlegenden Beiträge zur Koordinationsfunktion sowie Pfaff/Pfeiffer (2004) zu einem systematischen Überblick.

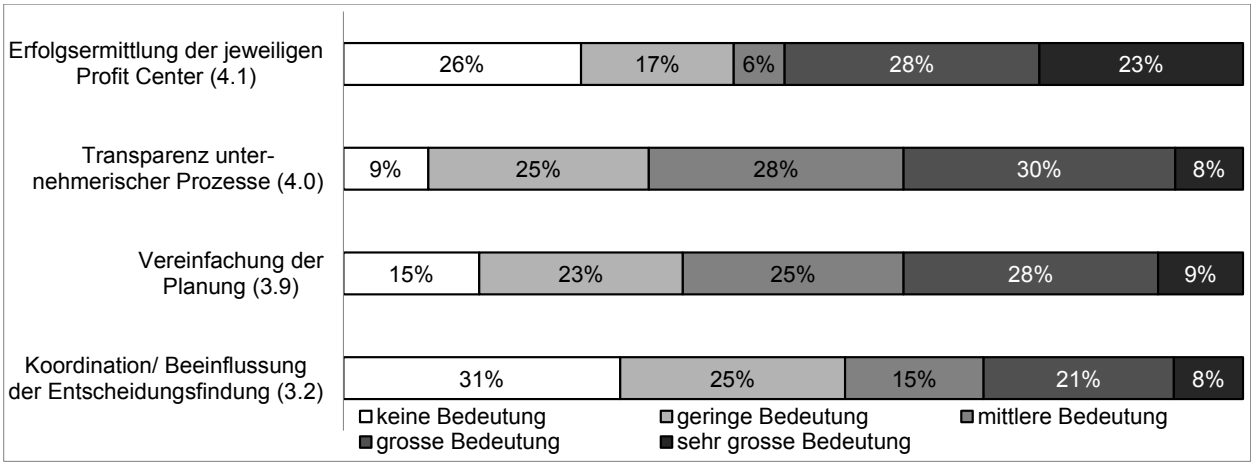


Abbildung 3: Bedeutung interner Funktionen der Verrechnungspreisgestaltung[21]

Bereits 2005 war die Erfolgsermittlung mit einem mittleren Wert von 3.54 (gemessen auf einer Skala von 1 «keine Bedeutung» bis 5 «sehr grosse Bedeutung») die wichtigste Funktion, gefolgt von der Koordinationsfunktion mit 2.9. Allerdings bezogen sich die damals gemessenen Werte nur auf den Fall des Lieferungs- oder Leistungsaustauschs innerhalb eines Steuersubjekts, bei der die Steuerproblematik grundsätzlich zu vernachlässigen ist. Umso bemerkenswerter ist die relativ hohe Bedeutung der Funktionen Erfolgsermittlung und Transparenz in der aktuellen Umfrage einzuschätzen.

Abbildung 4 zeigt die Wichtigkeit externer Funktionen der Verrechnungspreisgestaltung. Der steuerlichen Compliance (Erfüllung steuerrechtlicher Regelungen) wird mit einer durchschnittlichen Bewertung von 5.6 die grösste Bedeutung sämtlicher Funktionen (interner wie externer) beigemessen. Diese Beobachtung spiegelt die Komplexität steuerrechtlicher Regelungen sowie die Entschlossenheit ausländischer Fisci bei der Durchsetzung ihrer Verrechnungspreisrichtlinien wider. Unternehmen erachten es als oberste Priorität, dass ihre Verrechnungspreise von den betroffenen Steuerbehörden akzeptiert werden. Dies deckt sich mit den Ergebnissen von Hummel für Unternehmen mit Sitz in Deutschland.[22]

Die zweitgrösste Bedeutung der externen Funktionen wird mit einem Durchschnittswert von 4.5 der Optimierung der gesamten Konzernsteuerlast beigemessen. Dies verdeutlicht, dass Konzerne innerhalb der Grenzen internationaler Verrechnungspreisstandards und -richtlinien nach wie vor versuchen, die Steuerbelastung der verbundenen Unternehmen zu optimieren. Auch in der Umfrage 2005 zeigte sich die Optimierung der Konzernsteuerlast als wichtige Funktion, auch wenn die Ergebnisse nicht unmittelbar vergleichbar sind, da die steuerliche Compliance als Antwortalternative gefehlt hatte. Weitere externe Funktionen wie die handelsrechtliche Bestandsbewertung, die Segmentberichterstattung oder die Preisbestimmung auf Absatzmärkten sind von vergleichsweise geringer Bedeutung.

21 Pfaff/Hummel (2014), S. 593.

22 Vgl. Hummel (2010), S. 159.

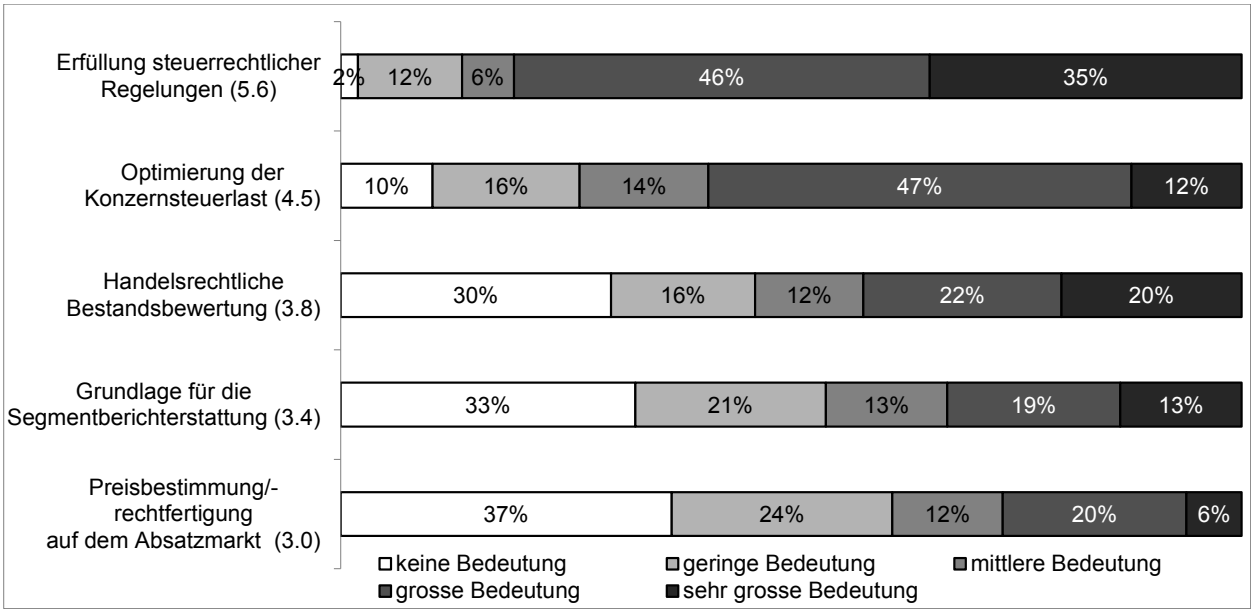


Abbildung. 4: Bedeutung externer Funktionen der Verrechnungspreisgestaltung[23]

Zusammenfassend kann festgehalten werden, dass steuerliche Überlegungen die Verrechnungspreisgestaltung der befragten Schweizer Unternehmen dominieren. Dabei liegt das Hauptaugenmerk auf der steuerlichen Compliance und weniger auf der Optimierung der Konzernsteuerlast, was im Einklang mit empirischen Befunden zur zunehmenden Bedeutung des Tax Risk Managements bei international tätigen Unternehmen steht. Einer Befragung von Ernst & Young zufolge gewichten global agierende Öl- und Gasunternehmen das Tax Risk Management im Jahre 2013 deutlich stärker als in den Jahren 2010 und 2007; damit einhergehend sinken Bestrebungen zur Steueroptimierung bei diesen Unternehmen.[24]

2.1.4. Methoden der Verrechnungspreisgestaltung

2.1.4.1. Unterscheidung nach der Art der Transaktionsbeziehung

Massstab für die steuerliche Angemessenheit der Verrechnungspreisgestaltung ist international der Fremdvergleichsgrundsatz, der sich auf «die wirtschaftlich relevanten Eigenschaften der zu vergleichenden Umstände» bezieht.[25] Vergleichbarkeit heisst, «dass keiner der Unterschiede (soweit vorhanden) zwischen den zu vergleichenden Umständen die untersuchte Bedingung der Methode (z.B. Preis oder Spanne) wesentlich beeinflussen kann oder dass hinreichend genaue Berichtigungen erfolgen können, um die Auswirkung dieser Unterschiede zu beseitigen.»[26] Darauf gestützt beschreibt die Verrechnungspreisrichtlinie der OECD eine Reihe von Methoden, «die verwendet werden können, um festzustellen, ob die in den kaufmännischen oder finanziellen Beziehungen zwischen verbundenen

23 Pfaff/Hummel (2014), S. 594.

24 Vgl. Ernst & Young (2013).

25 OECD (2011a), S. 45.

26 OECD (2011a), S. 45; zu ausführlichen Leitlinien zur Durchführung von Vergleichbarkeitsanalysen siehe OECD (2011c).

Unternehmen festgelegten Bedingungen mit dem Fremdvergleichsgrundsatz vereinbar sind.»[27] Ziel der Richtlinie ist es, dass die für die Umstände des Einzelfalls am besten geeignete Methode ausgewählt wird.[28] Dabei erweist sich eine Unterscheidung nach der Art der Lieferungs- und Leistungsbeziehungen als sinnvoll. Im Fragebogen wurden vereinfachend folgende drei grundlegende Typen von Geschäftsvorfällen unterschieden:

- Transaktionen zwischen Produktionsgesellschaften,
- Transaktionen zwischen Produktions- und Vertriebsgesellschaft sowie
- Erbringung zentraler Dienstleistungen.

Als Methoden zur internen Lieferungs- und Leistungsverrechnung lassen sich – zunächst unabhängig von der Art des Geschäftsvorfalls – die von der OECD vorgeschlagenen Standardmethoden (Preisvergleichs-, Wiederverkaufspreis- sowie Kostenaufschlagsmethode) und Gewinnmethoden (Nettomargen- sowie Gewinnaufteilungsmethode) unterscheiden. Im Folgenden werden die genannten Methoden kurz erläutert. Nicht auf dem Fremdvergleichsgrundsatz beruhende Ansätze wie namentlich die globale formelhafte Aufteilung des Gewinns werden von der OECD (bislang) vehement abgelehnt[29] und bleiben in der Analyse unberücksichtigt.

Gemäss OECD-Verrechnungspreisrichtlinien 2010 vergleicht die Preisvergleichsmethode «den bei einem konzerninternen Geschäftsvorfall verrechneten Waren- oder Dienstleistungspreis mit jenem, der bei einem vergleichbaren Geschäftsvorfall zwischen unabhängigen Unternehmen unter vergleichbaren Verhältnissen verrechnet wird.»[30] Die *Preisvergleichsmethode* wird dann als besonders geeignete Methode angenommen, wenn «ein unabhängiges Unternehmen dasselbe Produkt verkauft, das auch zwischen zwei verbundenen Unternehmen verkauft wird.»[31] Preisanpassungen sind erforderlich, wenn sich z.B. die Lieferbedingungen hinsichtlich Transport und Versicherung unterscheiden.[32]

Die *Wiederverkaufspreismethode* «geht von dem Preis aus, zu dem ein Produkt, das von einem verbundenen Unternehmen gekauft worden ist, an ein unabhängiges Unternehmen weiterveräußert wird.»[33] Dieser um eine angemessene Bruttomarge (Handelsspanne) sowie um Drittkosten, die mit dem Kauf in Zusammenhang stehen (z.B. Transport- und Zollkosten), berichtigte Preis ergibt den Verrechnungspreis für den konzerninternen Verkauf. Die Bruttomarge umfasst die Vertriebs- und sonstigen betrieblichen Kosten, die der Wiederverkäufer (des gruppenintern bezogenen Produkts) zu decken hat, sowie den unter Berücksichtigung der wahrgenommenen Funktionen zu erzielenden angemessenen Gewinn.[34] Die Wiederverkaufspreismethode eignet sich vor allem dann, wenn die

27 OECD (2011b), S. 65.

28 Vgl. OECD (2011b), S. 65 ff.

29 Vgl. OECD (2011a), S. 39-44.

30 OECD (2011b), S. 70.

31 OECD (2011b), S. 71.

32 Vgl. OECD (2011b), S. 72.

33 OECD (2011b), S. 72.

34 Vgl. OECD (2011b), S. 72 f.

empfangende Division das gelieferte Produkt direkt an Dritte weiterverkauft. Häufig handelt es sich dabei um Transfers zwischen einer Produktions- und einer Vertriebsgesellschaft.[35]

Bei der *Kostenaufschlagsmethode* wird der Verrechnungspreis (Fremdvergleichspreis) bestimmt, indem auf die Kosten des liefernden (oder leistenden) Bereichs ein angemessener (betriebs- oder branchenüblicher) Kostenzuschlag erhoben wird, der einen den wahrgenommenen Funktionen und Marktbedingungen entsprechenden Gewinn gewährleisten soll.[36] Gemäss OECD dürfte die Methode dann besonders zweckmässig sein, «wenn zwischen verbundenen Unternehmen Halbfabrikate verkauft werden, wenn verbundene Unternehmen Verträge über die Nutzung gemeinsamer Einrichtungen bzw. langfristige Abnahme- und Liefervereinbarungen abschließen oder wenn es beim konzerninternen Geschäftsvorfall um die Erbringung von Dienstleistungen geht.»[37]

Die Funktionsweise der *transaktionsbezogenen Nettomargenmethode* ist ähnlich wie die der beiden zuvor genannten Methoden. Untersucht wird der Nettogewinn, den ein verbundenes Unternehmen aus einem konzerninternen Geschäft erzielt, «in Relation zu einer geeigneten Grundlage (z.B. Kosten, Umsatz, Kapital).»[38] Dabei entscheidend ist die Wahl geeigneter Vergleichsunternehmen (Funktion, Risikostruktur etc.). Ob Vergleichbarkeit vorliegt, soll anhand einer Funktionsanalyse der Geschäftsvorfälle zwischen den verbundenen Unternehmen sowie zwischen unabhängigen Dritten geprüft werden.[39] Die Nettomargenmethode gilt als unzuverlässig, «wenn jede Partei wertvolle, einzigartige Beiträge leistet.»[40] In diesem Fall wird auf die Gewinnaufteilungsmethode verwiesen.

Die *Gewinnaufteilung* wird danach bestimmt, was unabhängige Unternehmen bei einem gegebenen Geschäftsvorfall erwartet hätten. Dabei wird zunächst der aus dem Geschäftsvorfall (der Geschäftsbeziehung) erwirtschaftete Gesamtgewinn ermittelt und dieser dann nach wirtschaftlich angemessenen Kriterien aufgeteilt; die Aufteilung soll dem nahekommen, «was bei einem zwischen fremden Dritten vereinbarten Geschäftsvorfall erwartet worden wäre.»[41] Vorteile der Methode werden vor allem dann gesehen, wenn

- Transaktionen hochintegrierte Tätigkeiten umfassen (wie z.B. den weltweiten Handel mit Finanzinstrumenten zwischen verbundenen Unternehmen) oder
- die verbundenen Unternehmen einzigartige und wertvolle Beiträge zum gemeinsamen Geschäft leisten.[42]

2.1.4.2. Verrechnungspreisgestaltung zwischen Produktionsgesellschaften

Bei Transaktionen zwischen Produktionsgesellschaften wird typischerweise ein Zwischen- oder Halbfabrikat zur Weiterverarbeitung oder Modifizierung an ein verbundenes Unternehmen verkauft. Das

35 Vgl. OECD (2011b), S. 73.

36 Vgl. OECD (2011b), S. 78 f.

37 OECD (2011b), S. 79.

38 OECD (2011b), S. 86.

39 Vgl. OECD (2011b), S. 87.

40 OECD (2011b), S. 87.

41 OECD (2011b), S. 105.

42 Vgl. OECD (2011b), S. 105.

gelieferte Produkt wird nach Erhalt also nicht direkt abgesetzt, sondern fliesst in den Produktionsprozess des empfangenden Unternehmens ein.

Abbildung 5 verdeutlicht, welcher Anteil der Lieferungen zwischen Produktionsgesellschaften (bzw. zwischen Produktions- und konzerninternen Distributionsgesellschaften) mit der jeweiligen Verrechnungspreismethode bewertet wird.[43] Varianten der Kostenaufschlagsmethode werden durchschnittlich für 43 Prozent aller Transfers zwischen Produktionsgesellschaften verwendet, wobei die befragten Unternehmen erwartungsgemäss deutlich öfter einen Aufschlag auf die vollen als auf die variablen Kosten vornehmen. Daneben kommt die Wiederverkaufspreismethode am häufigsten zur Anwendung. Im Durchschnitt werden 18 Prozent der transferierten Zwischen- oder Halbfabrikate anhand dieser Methode bewertet, während Vergleiche mit Marktpreisen (Preisvergleichsmethode) bei rund 12 Prozent der erzielten Umsätze angestellt werden. Ebenfalls nennenswert ist die (transaktionsbezogene) Nettomargenmethode, welche bei nahezu 9 Prozent aller Transaktionen Anwendung findet. Die (geschäftsvorfallbezogene) Gewinnaufteilungsmethode, aber auch weitere gewinnorientierte Methoden – wie der globale Gewinnvergleich sowie die globale Gewinnaufteilung – sind für die befragten Unternehmen praktisch bedeutungslos.

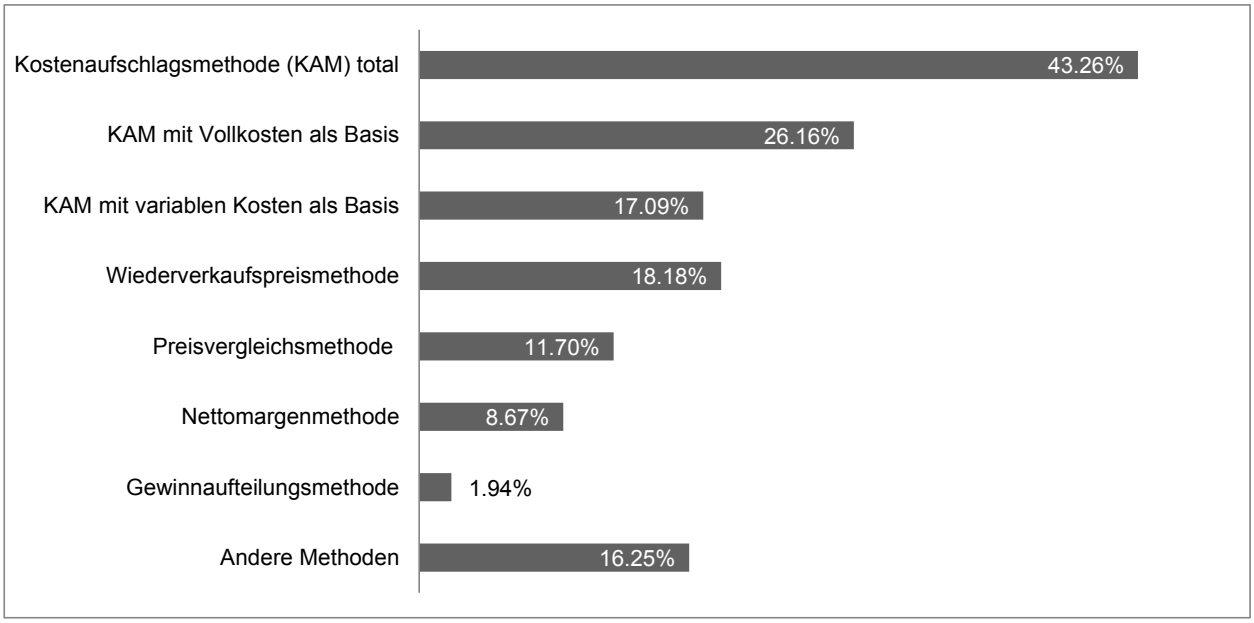


Abbildung 5: Methoden der Verrechnungspreisgestaltung zwischen Produktionsgesellschaften (bzw. zwischen Produktions- und konzerninternen Distributionsgesellschaften)

Insgesamt bestätigt sich die Empfehlung der OECD, dass bei Verbundbeziehungen zwischen Produktionsgesellschaften die Kostenaufschlagsmethode am zuverlässigsten sein dürfte. Addiert man die Unternehmen, die rein zu variablen oder vollen Kosten ohne Aufschlag abrechnen, hinzu (in der Grafik unter «Andere Methoden» subsummiert), bestimmt gut die Hälfte der befragten Unternehmen den Verrechnungspreis kostenorientiert. Gleichzeitig zeigt sich, dass aber auch andere von der OECD vorgeschlagene Methoden – insbesondere die Preisvergleichs- und Wiederverkaufspreismethode – Anwendung finden.

43 In Fällen, in denen die Summe sämtlicher deklarierter Anteile eines Unternehmens nicht 100 Prozent entsprach, wurden die Angaben unter Beibehaltung der relativen Anteile reskaliert, so dass die Summe 100 Prozent ergibt.

Ein Vergleich zu den Umfrageergebnissen 2005 findet geschäftsfallübergreifend am Ende von Kapitel 4 statt; die damalige Umfrage hatte noch nicht nach Geschäftsvorfällen getrennt.

2.1.4.3. Verrechnungspreisgestaltung zwischen Produktions- und Vertriebsgesellschaften

Transaktionen zwischen Produktions- (bzw. Distributions-) und Vertriebsgesellschaften liegen vor, wenn die liefernde Gesellschaft ein zum Verkauf fertiges Produkt an den Vertrieb liefert, das dieser in der Folge an unabhängige Dritte absetzt. *Abbildung* 6 veranschaulicht die Bedeutung der verschiedenen Verrechnungspreismethoden für diesen Geschäfts(vor)fall. Im Vergleich zur Lieferung zwischen Produktionsgesellschaften gewinnt nun die von der OECD für diesen Typ empfohlene Wiederverkaufspreismethode mit rund einem Drittel aller Transaktionen stark an Bedeutung. Dennoch stellt die Kostenaufschlagsmethode immer noch die am häufigsten eingesetzte Verrechnungspreismethode dar: Auch zwischen Produktions- und Vertriebsgesellschaften werden noch knapp 36 Prozent aller Transfers mittels Kostenaufschlagsmethode bewertet. Die Preisvergleichsmethode sowie die transaktionsbezogene Nettomargenmethode kommen mit knapp 10 Prozent etwa gleich stark wie im ersten Fall zum Einsatz. Alle anderen Methoden sind eher unbedeutend.

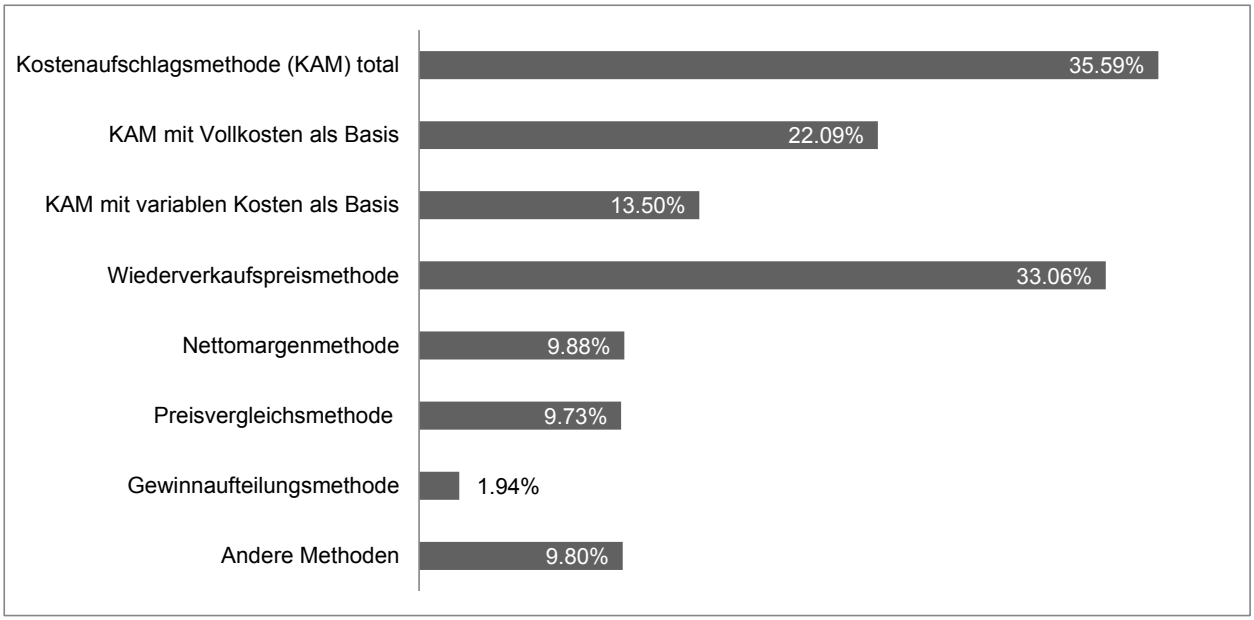


Abbildung 6: Methoden der Verrechnungspreisgestaltung zwischen Produktions- (bzw. Distributions-) und Vertriebsgesellschaften

2.1.4.4. Verrechnungspreisgestaltung für zentrale Dienstleistungen

Konzerne stellen ihren verbundenen Unternehmen typischerweise eine Reihe von Dienstleistungen zur Verfügung. Dazu gehören administrative und kaufmännische Dienstleistungen (inklusive Management-, Koordinations- und Kontrollfunktionen) ebenso wie Finanzdienstleistungen und Dienstleistungen aus dem Forschungs- und Dienstleistungsbereich. Die Erbringung von Dienstleistungen kann auch mit der

Übertragung von Waren oder immateriellen Wirtschaftsgütern verbunden sein.[44] Bei der Analyse der Verrechnungspreise muss geprüft werden, ob

- die konzerninterne Dienstleistung überhaupt erbracht worden ist und
- der Fremdvergleichsgrundsatz eingehalten wird.[45]

Als fremdvergleichskonforme Verrechnungspreismethoden werden von der OECD namentlich die Preisvergleichsmethode sowie die Kostenaufschlagsmethode erachtet: «Die Preisvergleichsmethode wird wahrscheinlich die zweckmäßigste Methode sein, wenn eine vergleichbare Dienstleistung zwischen unabhängigen Unternehmen auf dem Markt des Empfängers erbracht wird oder vom verbundenen Unternehmen, das die Dienstleistungen unter vergleichbaren Verhältnissen für ein unabhängiges Unternehmen erbringt. Dies trifft beispielsweise dann zu, wenn Dienstleistungen im Bereich des Rechnungswesens und der Revision sowie juristische Dienstleistungen oder Computerdienstleistungen erbracht werden, vorausgesetzt die konzerninternen Geschäftsvorfälle und die Fremdgeschäftsvorfälle sind vergleichbar. Die Kostenaufschlagsmethode ist in Ermangelung einer Preisvergleichsmethode wahrscheinlich dann die zweckmäßigste Methode, wenn die Art der Tätigkeit, der Kapitaleinsatz und die übernommenen Risiken mit denjenigen von unabhängigen Unternehmen vergleichbar sind.» [46]

Abbildung 7 zeigt, dass die Preisvergleichsmethode nur bei etwa 8 Prozent der befragten Unternehmen zum Einsatz kommt. Offensichtlich sind die konzernintern erbrachten Dienstleistungen derart individuell, dass ein Preisvergleich unzweckmässig wäre. Auch die Wiederverkaufspreismethode kommt mit 2.8 Prozent sehr selten zum Einsatz.

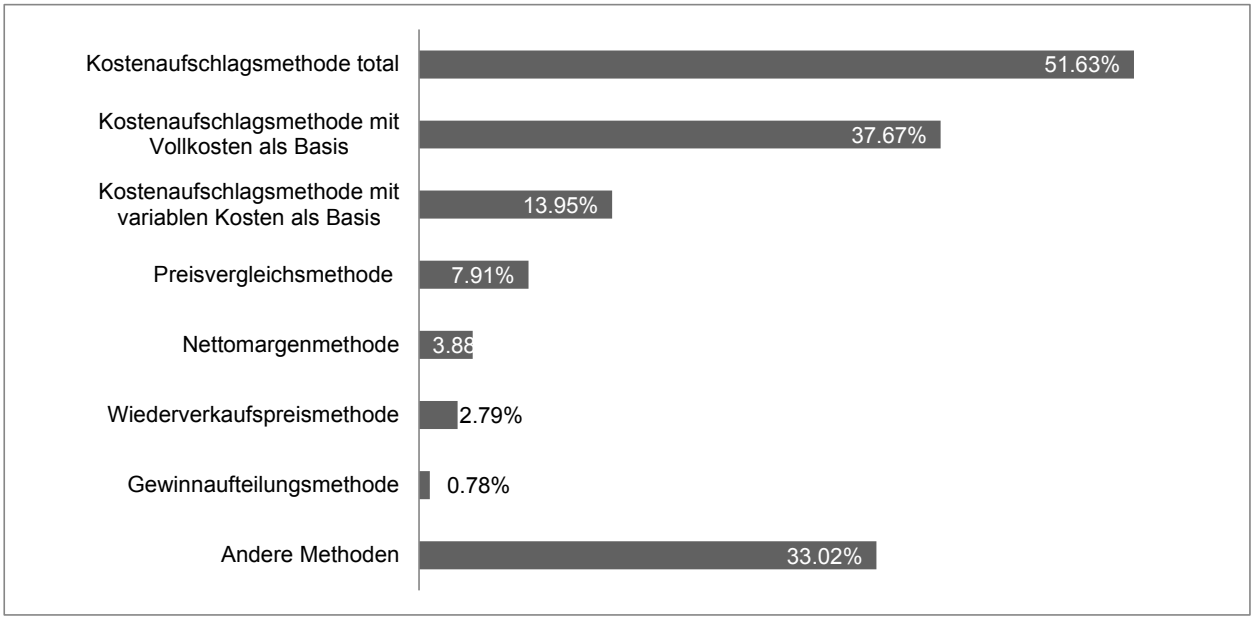


Abbildung 7: Methoden der Verrechnungspreisgestaltung für zentrale Dienstleistungen

Wenig überraschend dominieren daher kostenorientierte Verrechnungspreise. Rund 38 Prozent (bzw. 14 Prozent) der intern erbrachten Dienstleistungen werden durch einen Aufschlag auf die vollen (bzw.

44 Vgl. OECD (2011d), S. 233 f.

45 Vgl. OECD (2011d), S. 234 f.

46 OECD (2011d), S. 243 f.

variablen) Kosten verrechnet, zudem 15 Prozent (bzw. 2 Prozent) zu vollen (bzw. variablen) Kosten ohne Aufschlag. Des Weiteren werden durchschnittlich 7 Prozent der Dienstleistungen mittels Kostenumlage nach dem Pool-Konzept verrechnet. Insgesamt werden somit durchschnittlich 76 Prozent der zentral erbrachten Dienstleistungen kostenbasiert verrechnet. Es kann vermutet werden, dass kostenorientierte Verrechnungspreismethoden für konzerninterne Dienstleistungen einfacher anwendbar und dokumentierbar sind. Alle anderen Methoden, namentlich die Gewinnaufteilungsmethode sowie die transaktionsbezogene Nettomargenmethode, aber auch die Preisvergleichsmethode spielen bei der Verrechnung konzerninterner Dienstleistungen eine untergeordnete Rolle.

2.1.4.5. Vergleich mit den Umfrageergebnissen 2005[47]

Einschränkend muss zunächst festgestellt werden, dass die Ergebnisse 2012 und 2005 nicht unmittelbar miteinander vergleichbar sind, da im Fragebogen 2005 die Art der Transaktionsbeziehung noch keine Rolle spielte und die Methoden unterschiedlich gegliedert wurden. Bringt man aber die Daten 2012 und 2005 auf einen gemeinsamen Nenner, lassen sich zumindest die folgenden wichtigen Aussagen treffen:

Kosten- und marktpreisorientierte Methoden dominieren nach wie vor die Gestaltung der Verrechnungspreise der befragten Unternehmen. Bei den kostenorientierten Methoden liegt die Kostenaufschlagsmethode bei Zugrundelegung der vollen Kosten deutlich vorn. Die Verrechnung reiner Kosten (ohne Aufschlag) spielt eine untergeordnete Rolle. Dies gilt letztlich auch für alle Formen der Gewinnaufteilung sowie für die transaktionsbezogene Nettomargenmethode. Trotz Wegfall der OECD-Hierarchie der Verrechnungspreismethoden im Jahr 2010 hat sich an der Beliebtheit der Standardmethoden offensichtlich wenig geändert. Es kann nur spekuliert werden, dass gerade Kosten und Preise *die* beobachtbar harten Fakten sind, an denen sich Unternehmen bei der Anwendung des Fremdvergleichsgrundsatzes vorzugsweise orientieren.

2.1.5. Besondere Herausforderungen und Zufriedenheit der Unternehmen mit ihrer Verrechnungspreisgestaltung

Grundsätzlich können Konflikte zwischen allen in Kapitel 3 genannten Funktionen auftreten. Am wenigsten wahrscheinlich sind aus wissenschaftlicher Sicht Konflikte zwischen (interner) Erfolgsermittlung und steuerlicher Compliance: Da Verrechnungspreise danach zu beurteilen sind, ob sich verbundene Unternehmen wie unabhängige Fremde verhalten, wird eine «Gewinnmaximierungsabsicht der verbundenen Unternehmen unterstellt, die sich komplementär zur Erfolgsermittlungsfunktion verhält.»[48] Vor dem Hintergrund dieser Argumentation wird man sogar eine komplementäre Beziehung erwarten können. Diese Erwartung lässt sich empirisch nicht bestätigen. Im Mittelwert wird mit 3.6 sogar eine schwach konfliktäre Beziehung gesehen. *Abbildung* 8 zeigt zudem, dass die wahrgenommene Bedeutung stark streut: Sie reicht von der Einschätzung «keine Bedeutung» bis hin zu einer sehr grossen Bedeutung. Am stärksten wird mit einem Mittelwert von 4.1 der Konflikt zwischen (interner) Erfolgsermittlung und Steueroptimierung wahrgenommen.

47 Vgl. Pfaff/Stefani (2007).

48 Pfaff/Hummel (2014), S. 597.

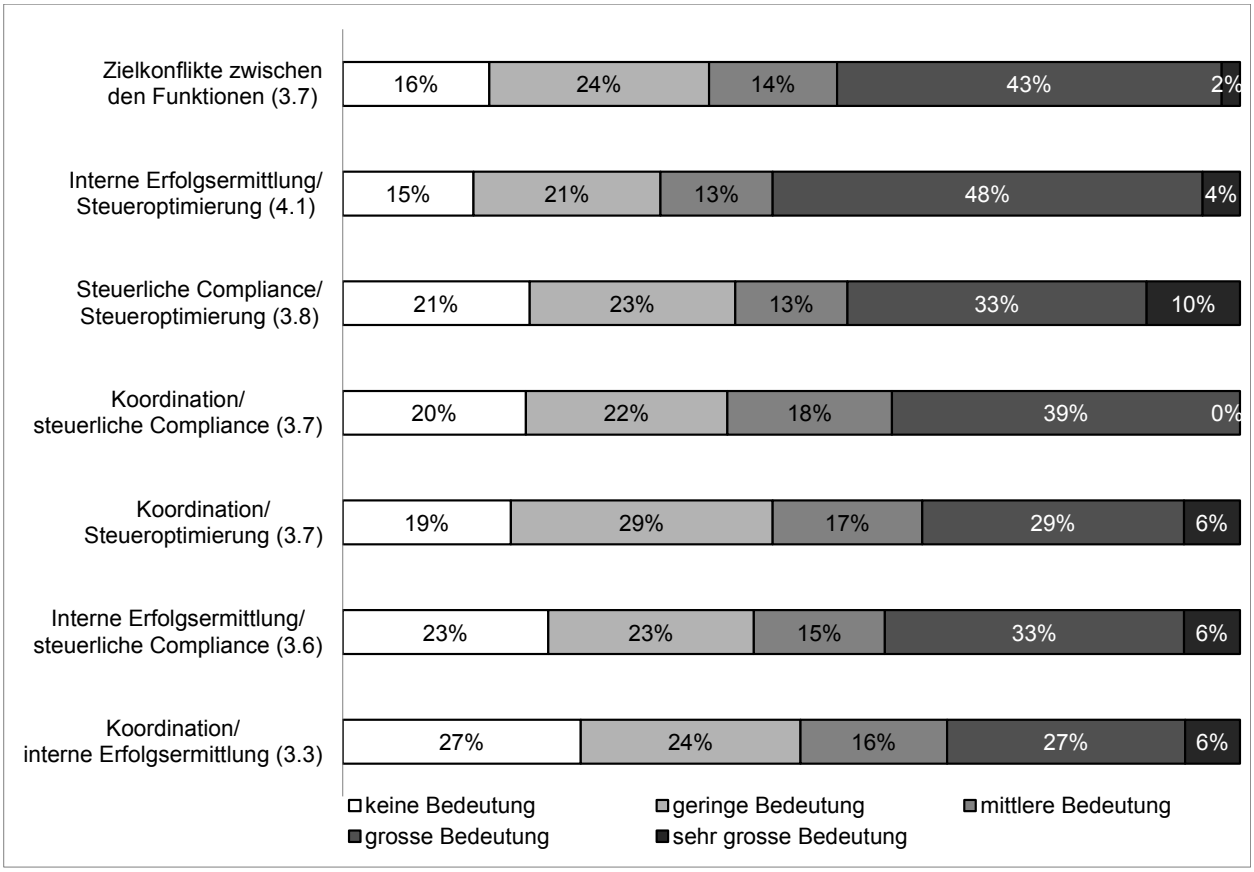


Abbildung 8: Bedeutung von Zielkonflikten zwischen den Funktionen der Verrechnungspreisgestaltung[49]

Insgesamt ist festzustellen, dass Zielkonflikte zwischen den Verrechnungspreisfunktionen für die befragten Unternehmen offensichtlich keine sehr grosse Herausforderung darstellen (durchschnittliche Bewertung von 3.7 auf einer Skala von 1 bis 7). Dieser Befund wird dadurch gestützt, dass Unternehmen bei Anwendung der Kostenaufschlagsmethode kaum mit Transparenznachteilen rechnen. In der Literatur wird demgegenüber befürchtet, dass die Kostenaufschlagsmethode bei mehrstufiger Wertschöpfungskette zu einer Mehrfachbezuschlagung und damit zu einer Verschleierung der tatsächlichen Kostenstruktur führen könnte.[50] Die Koordinationsfunktion, welche die aus Konzernsicht gewinnmaximierenden Entscheidungen im Fokus hat, könnte dann in Konflikt zur steuerlichen Compliance stehen, welche insbesondere bei mehrstufiger Wertschöpfungskette die Kostenaufschlagsmethode erfordern könnte. Fragt man die Unternehmen konkret nach diesen Gefahren, verweisen sie mit einer hohen durchschnittlichen Bewertung in Höhe von 5.2 (auf einer Skala von 1 bis 7) darauf, dass in derartigen Situationen auf Informationen des Kostenrechnungssystems zurückgegriffen wird.

Dies deutet darauf hin, dass das Verrechnungspreissystem von den befragten Unternehmen keineswegs isoliert gesehen wird, sondern vielmehr in das Management Control System (MCS) integriert ist (vgl. *Abbildung* 9).

49 Pfaff/Hummel (2014), S. 598.

50 Vgl. Pfaff/Hummel (2014), S. 598; Pfaff (2003), S. 37-40.

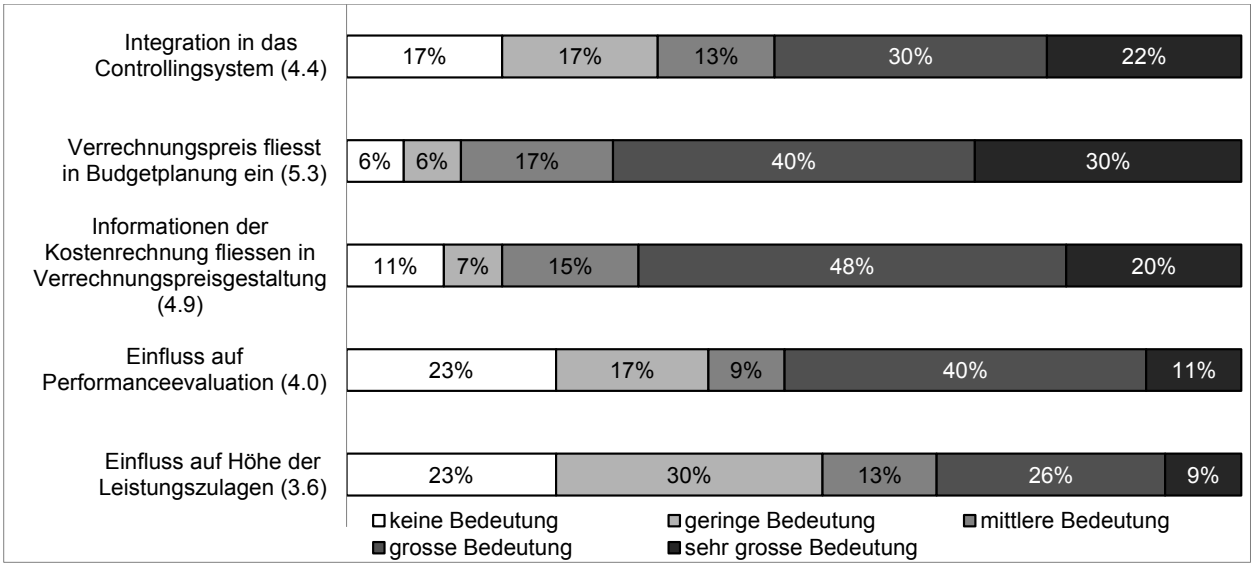


Abbildung 9: Integration von Verrechnungspreissystem und Management Control System

Diese Integration wird von den Befragungsteilnehmern insgesamt als relativ wichtig bewertet (durchschnittliche Bewertung von 4.4 auf einer Skala von 1 bis 7). Wie zu erwarten ist, findet die Integration insbesondere im Rahmen der Budgetplanung, aber auch im Hinblick auf die Performanceevaluation der Leiter der Verantwortungsbereiche statt. Darüber hinaus fliessen Informationen des Kostenrechnungssystems direkt in die Verrechnungspreisgestaltung ein, was im Einklang mit der häufigen Verwendung kostenbasierter Verrechnungspreismethoden steht (vgl. Abschnitte 4.2 bis 4.4).

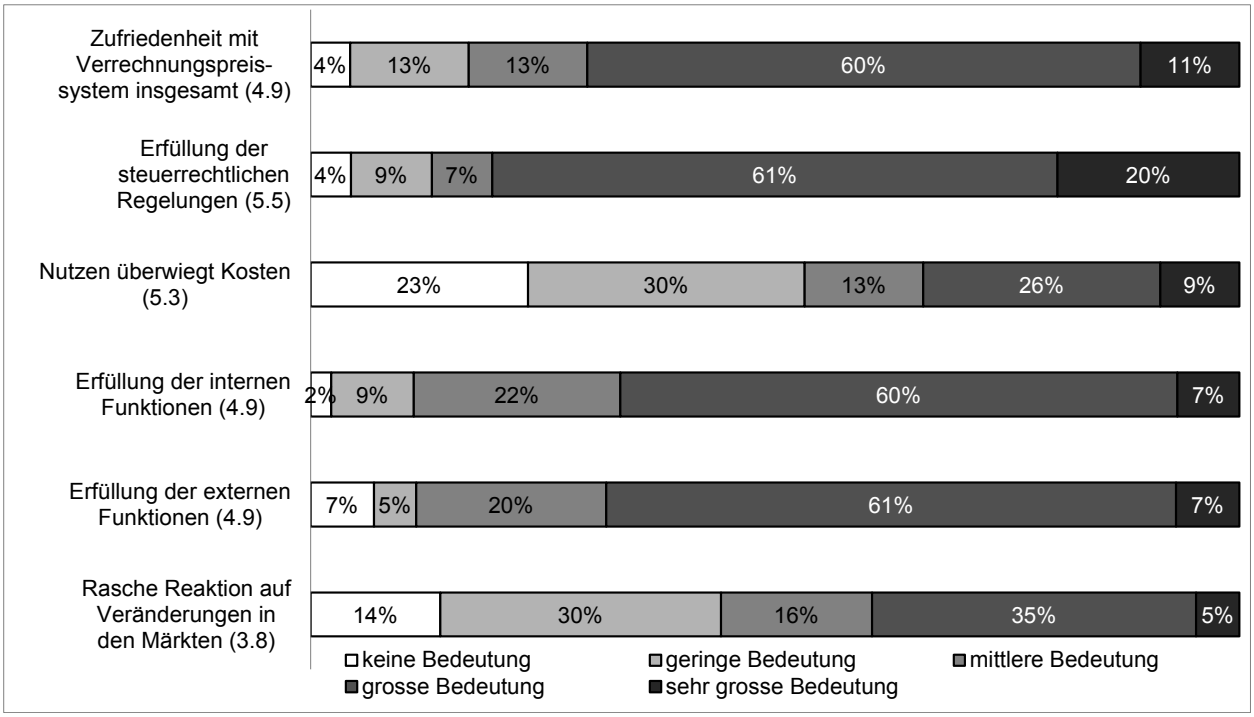


Abbildung 10: Zufriedenheit mit dem Verrechnungspreissystem

Zum Gesamtbild passt die im Durchschnitt hohe Zufriedenheit der befragten Unternehmen mit ihrem Verrechnungspreissystem (vgl. *Abbildung* 10). Während der Gesamtzufriedenheitswert im Durchschnitt

bei 4.9 (auf einer Skala von 1 bis 7) liegt, bewerten die Unternehmen sogar mit einem Wert von 5.5 (ebenfalls auf einer Skala von 1 bis 7) die Erfüllung steuerrechtlicher Regelungen. Diese externe Funktion der Verrechnungspreisgestaltung, der die befragten Unternehmen eine sehr hohe Bedeutung beimessen (vgl. Abschnitt 3), scheinen die in der Praxis implementierten Verrechnungspreissysteme somit weitestgehend zu erfüllen. Vor diesem Hintergrund ist auch die sehr positive Bewertung des Kosten-Nutzen-Verhältnisses der implementierten Verrechnungspreissysteme nicht überraschend: Mit einem Wert von 5.3 (ebenfalls auf einer Skala von 1 bis 7) stimmen die Unternehmen zu, dass der Nutzen des Verrechnungspreissystems im Grossen und Ganzen die Kosten überwiegt. Am wenigsten zufrieden zeigen sich die Unternehmen mit der Anpassungsfähigkeit ihrer Verrechnungspreissysteme im Hinblick auf Veränderungen in den Beschaffungs- und Absatzmärkten (durchschnittliche Bewertung von 3.8).

Weiterführende ökonometrische Methoden zeigen zudem, dass die Zufriedenheit der befragten Unternehmen mit der Integration des Verrechnungspreissystems korreliert ist. So steigt die Zufriedenheit statistisch signifikant mit der (angegebenen) Stärke der Integration. Zudem zeigt sich, dass die Transparenz des Verrechnungspreissystems gegenüber den beteiligten Parteien von grosser Wichtigkeit ist und ebenfalls mit der Integration sowie Zufriedenheit zusammenhängt. Einmal mehr scheint sich die Weisheit zu bewahrheiten, dass Kommunikation alles ist.

2.1.6. Thesenförmige Zusammenfassung

- Verrechnungspreise gehören nach wie vor zu den wichtigsten und meistdiskutierten Instrumenten dezentraler Steuerung. Grund dafür ist der systemimmanente Interessenkonflikt zwischen den beteiligten Profit Centern; bei grenzüberschreitenden Verrechnungspreisen treten die Partikularinteressen der Fisci sowie der Zollbehörden der betroffenen Länder hinzu.
- In dieser Situation gibt es weder wissenschaftlich eindeutig präferierte, noch einfache praktische Lösungen, da die aus der Kooperation der verbundenen Unternehmen entstehenden Synergien nur willkürlich aufgeteilt werden können.
- Trotz dieser schwierigen Ausgangslage sind Schweizer Konzerne mit der Gestaltung ihrer Verrechnungspreissysteme im Durchschnitt sehr zufrieden, so die Ergebnisse einer 2012 durchgeführten Umfrage des Lehrstuhls für Unternehmensrechnung und Controlling der Universität Zürich. Befragt wurden 158 Unternehmen, die Mitte Juli 2012 an der Schweizer Börse SIX primärkotiert waren. Die Rücklaufquote betrug 34 Prozent.
- Hinsichtlich der Funktionen gilt, dass steuerliche Überlegungen die Verrechnungspreisgestaltung der befragten Schweizer Unternehmen dominieren. Dabei liegt das Hauptaugenmerk auf der steuerlichen Compliance und weniger auf der Optimierung der Konzernsteuerlast. Zielkonflikte zwischen den Verrechnungspreisfunktionen stellen für die befragten Unternehmen offensichtlich keine grosse Herausforderung dar.
- Mit der Neufassung der OECD-Verrechnungspreisgrundsätze 2010 wurde die Dominanz der Standardmethoden aufgegeben. Neu soll stets die für jeden Fall am besten geeignete Methode gewählt werden. Die daraus abgeleitete Vermutung, dass die Standardmethoden (Preisvergleichsmethode, Wiederverkaufspreismethode, Kostenzuschlagsmethode) zu Gunsten

der Gewinnmethoden (namentlich die transaktionsbezogene Gewinnaufteilungsmethode sowie die transaktionsbezogene Nettomargenmethode) an Bedeutung verlieren, konnte nicht bestätigt werden.

- Unabhängig von der Transaktionsbeziehung (den beteiligten Wertschöpfungsstufen) weist die Kostenaufschlagsmethode wie bereits 2005 grundsätzlich die höchste Bedeutung auf. Am stärksten ist diese Dominanz bei der Verrechnung zentraler Dienstleistungen, am schwächsten bei Transaktionen zwischen Produktions- und Vertriebsgesellschaften eines Konzerns.
- Die Befürchtung, dass die Kostenaufschlagsmethode bei mehrstufiger Wertschöpfungskette zu einer Mehrfachbezuschlagung und damit zu einer Verschleierung der tatsächlichen Kostenstruktur führen könnte, wird von den befragten Unternehmen wenig geteilt. Vielmehr wird in diesen Situationen auf die Informationen des Kostenrechnungssystems zurückgegriffen.
- Mit der Zufriedenheit stark korreliert ist die Integration des Verrechnungspreissystems in das Management Control System sowie die Transparenz gegenüber den beteiligten Parteien. Einmal mehr scheint sich die Aussage zu bewahrheiten: «Kommunikation ist alles».

Literatur

Ernst & Young (2013): Navigating the choppy waters of international tax. 2013 Global Transfer Pricing Survey, www.ey.com/Publication/vwLUAssets/EY-2013_Global_Transfer_Pricing_Survey/$FILE/EY-2013-GTP-Survey.pdf.

Hirshleifer, J. (1956): On the Economics of Transfer Pricing. In: Journal of Business, 29. Jg., S. 172–184.

Hummel, K. (2010): Gestaltungsparameter und Einflussfaktoren von Verrechnungspreissystemen, Baden-Baden.

OECD (1995/96/97): Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations.

OECD (2011a): Der Fremdvergleichsgrundsatz. In: OECD-Verrechnungspreisrichtlinien für multinationale Unternehmen und Steuerverwaltungen 2010, OECD Publishing, S. 33-63.

OECD (2011b): Verrechnungspreismethoden. In: OECD-Verrechnungspreisrichtlinien für multinationale Unternehmen und Steuerverwaltungen 2010, OECD Publishing, S. 65-119.

OECD (2011c): Vergleichbarkeitsanalyse. In: OECD-Verrechnungspreisrichtlinien für multinationale Unternehmen und Steuerverwaltungen 2010, OECD Publishing, S. 121-147.

OECD (2011d): Besondere Überlegungen zu konzerninternen Dienstleistungen. In: OECD-Verrechnungspreisrichtlinien für multinationale Unternehmen und Steuerverwaltungen 2010, OECD Publishing, S. 233-248.

Peters, G./Pfaff, D. (2008/2011): Controlling. Wichtige Methoden und Techniken, 2. Aufl., Zürich.

Pfaff, D. (2003): Methodische Fragen einer internationalen Konzernkostenrechnung. In: Franz, K.-P./Hieronimus, A. (Hrsg.): Kostenrechnung im international vernetzten Konzern, Düsseldorf und Frankfurt, S. 29-46.

Pfaff, D./Hummel, K. (2014): Funktionen und Zielkonflikte der Verrechnungspreisgestaltung – Ergebnisse einer empirischen Erhebung in der Schweiz. In: Dobler, M. et al. (Hrsg.): Rechnungslegung, Prüfung und Unternehmensbewertung, Stuttgart, S. 587-604.

Pfaff, D./Pfeiffer, T. (2004): Verrechnungspreise und ihre formaltheoretische Analyse: Zum State of the Art. Ist das Dilemma der pretialen Lenkung wirklich ein Dilemma? In: Die Betriebswirtschaft, 64. Jg., S. 296–319.

Pfaff, D./Stefani, U. (2006): Verrechnungspreise in der Unternehmenspraxis. Eine Bestandsaufnahme zu Zwecken und Methoden. In: Controlling, 18. Jg., S. 517-524.

Pfaff, D./Stefani, U. (2007): Transferpreise in der Schweizer Unternehmenspraxis. Empirische Ergebnisse und betriebswirtschaftliche Steuerungsgefahren. In: Meyer, C./Pfaff, D. (Hrsg.): Jahrbuch Finanz- und Rechnungswesen 2007, Zürich, S. 199-225.

Ronen, J./McKinney, G. (1970): Transfer Pricing for Divisional Autonomy. In: Journal of Accounting Research, 8. Jg., S. 99–112.

Weilenmann, Paul (1989): Dezentrale Führung: Leistungsbeurteilung und Verrechnungspreise. In: Zeitschrift für Betriebswirtschaft, 59. Jg., S. 932-956.

Wolff, U. (2010): OECD: Bekanntgabe der überarbeiteten Verrechnungspreisgrundsätze (Kapitel I-III). In: Ernst & Young Tax News, Oktober, S. 3-4.

2.2. Can the Integration of a Tax Compliant Transfer Pricing System into the Management Control System Be Successful? Yes, It Can!

Abstract: This paper examines how the integration of a single-book tax compliant transfer pricing system into the management control system is related to the perceived success of that transfer pricing system. To identify the relationships, we use results from survey data from Swiss multinational firms with cross-border internal transactions. Whereas analytical research suggests a decoupling of transfer prices to attain both tax compliance and internal (control) purposes, we find that the integration of a tax compliant transfer pricing system into the management control system is perceived to be successful in achieving these goals. This is particularly true when the transfer pricing system is transparent and can be revised in the case of fundamental management control problems. Corroborating these findings, we also find affirmative testimonies from interview data from employees within three multinational enterprises. Overall, our results are consistent with survey-based findings reporting that firms predominantly use the same transfer price for both tax compliance and management control purposes.

Authors: Dr. Katrin Hummel, Prof. Dr. Dieter Pfaff, Benedikt Bisig

Keywords: transfer pricing system integration; enabling use; transfer pricing system success; survey and interview data

2.2.1. Introduction

As recent discussions about tax avoidance illustrate (Sikka & Willmott, 2010), transfer pricing in multinational enterprises with cross-border internal transactions is the cornerstone of the allocation of profits across different tax jurisdictions. To impede tax minimization strategies, international guidelines, such as the OECD transfer pricing guideline, govern the determination of transfer prices. As a consequence, transfer prices need to be compliant, in principle, with this regulation. At the same time, tax compliant transfer prices are usually to some extent integrated into the enterprises' management control systems through their use in the planning and control of decentralized responsibility centers and subsidiaries.

In light of numerous agency problems in firms, it has been determined that a single (tax compliant) transfer price cannot simultaneously fulfill both internal (management control) and external (tax compliance and tax burden minimization) objectives (Baldenius, Melumad, & Reichelstein, 2004; Choe & Hyde, 2007; Hyde & Choe, 2005; Smith, 2002a). This finding has also found its way into management accounting textbooks (e.g., Horngren, Datar, & Rajan, 2015, pp. 883-884; Kaplan & Atkinson, 2014, pp. 454, 463-464; Zimmermann, 2013, pp. 186, 191-198). In an attempt to provide solutions to these goal conflicts, researchers have repeatedly called for the use of different transfer prices for different purposes, in particular, for the use of one set of books for tax compliance/tax optimization and another for internal management purposes (Baldenius et al., 2004; Choe & Hyde, 2007).

Despite this overwhelming consensus among analytical researchers, it is yet not clear that a single (tax compliant) transfer price necessarily weakens the success of a transfer pricing system with respect to management control purposes. Survey-based findings report that firms predominantly use the same transfer price for both internal and external (tax compliance)

purposes (Ernst & Young, 2001, 2003, 2005). In addition, consultants emphasize that the use of the same transfer price for tax compliance and internal management purposes may be helpful to signal that the transfer pricing system is driven by internal control considerations and not by tax optimization purposes (Ernst & Young, 2001, 2003). Even analytical researchers acknowledge that multinational enterprises, in practice, use one set of books, "*both for simplicity and in order to avoid the possibility that multiple transfer prices become evidence in any disputes with the tax authorities*" (Baldenius et al., 2004, p. 592).

For these reasons, the advantages of decoupled over integrated tax compliant transfer prices are questionable. Furthermore, transfer pricing and management control systems are highly complex and dynamic instead of exogenously given and static (which are common assumptions in analytical modeling). In particular, responsibility centers are often able to initiate revisions of the transfer pricing system if fundamental problems occur and deviate from internal transfer pricing rules in some exceptional cases (Cools, Emmanuel, & Jorissen, 2008). However, to initiate such revisions, transparency is needed, and management must truly comprehend the nature of the system and the rationale behind the rules. Furthermore, responsibility center managers must understand the effect of their actions on the company as a whole to avoid taking actions that could harm the company.

Against this backdrop, the interplay between international transfer pricing and management control systems has recently gained attention in the transfer pricing literature (Cools, 2014; Cools et al., 2008; Cools & Slagmulder, 2009; Rossing & Rohde, 2010). These case-based studies analyze the consequences of implementing a single-book tax compliant transfer pricing system for the design and use of various components of the management control system. However, despite its importance for practitioners and academics alike, little empirical evidence

exists on how firms address the conflict between management control purposes and tax compliance and whether the *level* of integration of the transfer pricing system into the management control system is positively associated with the success of that transfer pricing system. Chen and Chen (2015) shed some light on the conflict with regard to the determinants and consequences of transfer pricing autonomy for divisional managers. However, to our knowledge, only Cools et al. (2008) provide insights into managers' perceptions of an integrated transfer pricing system. While one interviewee highlights the improved tax compliance of the integrated transfer pricing system, others criticize its lack of flexibility. Cools et al. (2008, p. 626) call for future research to further analyze the consequences of integrated transfer pricing systems. We respond to this call and investigate two related research questions using survey data: can (tax compliant) transfer pricing system integration be positively related to transfer pricing system success as perceived by the corporate-level authorities responsible for transfer pricing, and is the relationship mediated by the *use* of the tax compliant transfer pricing system in terms of repairability, transparency, and flexibility. To shed light on these questions, we conduct an empirical study that is characterized by the following aspects.

First, in contrast to previous studies on transfer pricing (for an overview, see Cools, 2014), this study employs a perspective that goes beyond the analysis of transfer pricing methods (e.g., cost plus or resale-minus transfer prices) applied in single transactions. In particular, we analyze a firm's transfer pricing system in its entirety. While our approach, to some extent, departs from common research ground, it enables us to account for the broad area of transfer price application within multinational enterprises because transfer pricing systems in practice often contain a variety of different transfer pricing methods for different transactions. Furthermore, taking a holistic view of taxpayers' intercompany transactions is in line with recent OECD

developments, such as the base erosion and profit shifting (BEPS) project, and it is a consequence of the changing view of tax authorities (Ernst & Young, 2013, p. 29).

Second, to measure the level of integration between the transfer pricing system and the management control system, we use information on how strictly transfer prices enter the budget planning, cost accounting, performance evaluation and bonus systems of the responsibility centers. All of these parts of the management control system are usually seen as instruments for achieving decentralization and coordination in multidivisional and multinational firms (see, for example, Horngren et al., 2015, p. 868; Kaplan & Atkinson, 2014; Zimmermann, 2013, pp. 185-188). Moreover, this type of measurement allows us to build on previous research in the field (Cools et al., 2008; Rossing & Rohde, 2010).

Third, with respect to the use of the transfer pricing system, we rely on Adler and Borys' (1996) concept of an enabling use of bureaucracy. This concept is well established in management accounting research (Ahrens & Chapman, 2004; Chapman & Kihn, 2009; Free, 2007; Wouters & Roijmans, 2011; Wouters & Wilderom, 2008) and is particularly useful for examining the simultaneous use of highly structured controls (Ahrens & Chapman, 2004, p. 276), a phenomenon that is also prevalent in transfer pricing.

Results from a partial least squares (PLS) analysis of survey data from 38 multinational enterprises indicate that the level of integration of the transfer pricing system into the management control system is positively and significantly correlated with the success of the transfer pricing system as perceived by the corporate-level authorities responsible for transfer pricing. Note that we measure transfer pricing system success based on assessments of the satisfaction with the transfer pricing system in terms of the fulfillment of both tax compliance

and management control objectives. A considerable part of this correlation is due to the mediating variables repair and internal transparency. More precisely, the level of transfer pricing system integration is positively associated with the ability to repair the transfer pricing system in the event of fundamental problems and with internal transparency, which in turn have positive links to transfer pricing system success. To substantiate our survey results in terms of a robustness test, we contrast them with findings based on interviews with the corporate-level managers responsible for transfer pricing at three multinational enterprises.

Our study makes several contributions to the literature. First, our study complements analytical research (Baldenius et al., 2004; Choe & Hyde, 2007; Hyde & Choe, 2005; Smith, 2002a, 2002b) by analyzing empirical data from a *diverse* set of firms and providing new insights into the question of how multinational enterprises handle the (supposed) trade-off between the tax compliance and management control objectives of transfer pricing. Second, prior research on the integration between transfer pricing and management control systems is exclusively based on case studies (Cools et al., 2008; Rossing & Rohde, 2010). Our cross-sectional analysis and additional interviews can help to substantiate these case-based findings and, in addition, provide reliable insights into different levels of integration, use and success of transfer pricing systems. Third, our perspective on transfer pricing, which is independent of transactions and transfer pricing methods, allows us to examine differences in the *use* of the transfer pricing system. Practical implications from our research include recommendations for the strong integration of tax compliant transfer prices into the management control system; further, this integration needs to be accompanied by internal transparency and the ability to repair the transfer pricing system. In other words, successful tax compliant transfer pricing systems are highly integrated into the management control system and enable responsibility centers both to deeply understand transfer

pricing determination and to initiate revisions of the transfer pricing system if fundamental problems occur.

The remainder of this article is structured as follows. The second section reviews and summarizes the related literature. The third section elaborates the formulation of our hypotheses. The research design and data are described in the fourth section. The fifth section presents descriptive statistics and results from structural equation modeling along with results from additional robustness analyses and interview findings. The final section concludes the paper.

2.2.2. Prior Research and Conceptual Underpinnings

2.2.2.1. Related Transfer Pricing Literature

The transfer pricing literature addresses both the management control role and the taxation issues of transfer prices. While the internal objectives of transfer pricing have traditionally been at the forefront of (analytical) researchers' thinking,[1] more recent transfer pricing studies analyze the design of transfer prices with respect to both internal *and* external (in particular tax-related) objectives in decentralized organizations (Baldenius et al., 2004; Choe & Hyde, 2007; Halperin & Srinidhi, 1991; Hyde & Choe, 2005; Narayanan & Smith, 2000). Most of these researchers either focus on a single transfer price, which would optimally balance the conflict between tax optimization and internal resource allocation (Baldenius et al., 2004), or they decouple the internally used transfer price from the "*arm's length price*" by using two sets of books (Baldenius et al., 2004; Choe & Hyde, 2007; Hyde & Choe, 2005). While findings from

[1] The "standard" transfer pricing model was introduced by Hirshleifer (1956). In this model, the transfer price that optimally solves the resource allocation objective derives from an optimization program at the corporate level and thus falls short in explaining why companies use transfer prices instead of simply assigning optimal production volumes to the divisions. Subsequent studies therefore incorporate information asymmetry (e.g., Banker & Datar, 1992; Ronen & Balachandran, 1988) and incomplete contracting (e.g., Edlin & Reichelstein, 1995). A detailed review of the analytical transfer pricing literature is provided by Göx and Schiller (2007).

analytical research have advanced our understanding on the trade-offs between different transfer pricing objectives, these models are yet not able to account for the complex and dynamic nature of both transfer pricing systems and management control systems in practice. In particular, these analytical models share the common notion of a static and exogenously given management control system, a focus on a particular transaction and a disregard for variation in the level of integration between the transfer pricing system and the management control system. Moreover, these models do not incorporate how a transfer pricing system is used (in terms of transparency, repair, and flexibility).

Aside from analytical research on transfer pricing, there is a considerable number of empirical studies on transfer pricing. In a recent review of the empirical transfer pricing literature, Cools (2014, p. 14) distinguishes three research streams: early studies on the management control issues of transfer pricing, tax accounting studies and studies on the relationship between tax compliant transfer pricing systems and the design and use of management control systems. In this literature review, we concentrate on studies that explicitly investigate the relationship between transfer pricing and management control systems. These studies typically investigate transfer pricing at the subunit level (Boyns, Edwards, & Emmanuel, 1999; Colbert & Spicer, 1995; Cools et al., 2008; Eccles, 1985; Rossing & Rohde, 2010; Van der Meer-Kooistra, 1994; Van Helden, Van der Meer-Kooistra, & Scapens, 2001). While early studies primarily concentrate on *domestic* transfer pricing and thus neglect the tax compliance role of *international* transfer pricing (Boyns et al., 1999; Colbert & Spicer, 1995; Van der Meer-Kooistra, 1994), more recent case-based research has begun to provide in-depth investigations into the relationship between a firm's international transfer pricing and its management control

system (Cools et al., 2008; Rossing & Rohde, 2010). These researchers show that a single-book tax compliant transfer pricing system interacts with the design and use of various components of the management control systems, namely the overhead cost allocation, the budgeting and the performance evaluation system.

More precisely, Cools et al. (2008) find that the implementation of a single-book tax compliant transfer pricing system results in a more coercive use of the management control system. While a higher level of transparency is achieved under the new transfer pricing system, losses in flexibility occur "due to the uniform transfer pricing policy, which need[s] to be consistently applied under all circumstances" (Cools et al., 2008, p. 625). Similarly, Rossing and Rohde (2010, p. 212) reveal "an increase in the formalization of services and a discontinuation of allocations by divisions to business units […] in order to enhance external acceptance of overhead cost allocation." In addition, based on internal accounting data from a large company, Bouwens and Steens (2016) show that although the use of (tax compliant) full-cost transfer pricing can send upstream production into a death spiral, the retention of the price can serve as a credible commitment device to motivate managers to reduce cost. This result again illustrates the complexity of transfer pricing systems in practice.

Taken together, findings from both case studies indicate close interactions between the tax compliant transfer pricing system and the management control system. We are therefore particularly interested in whether the level of integration of the transfer pricing system into the management control system is positively associated with the success of the transfer pricing system as perceived by the corporate-level authorities responsible for transfer pricing. In so doing, we go beyond the analysis of transfer pricing (methods) in distinct internal transactions and concentrate on a general perspective of the integration of the transfer pricing system into the

management control system. In addition, we investigate the use of the transfer pricing system – enabling versus coercive – which is delineated in the following subsection (section 2.2).

2.2.2.2. Enabling versus Coercive Use of the Transfer Pricing System

The distinction between an enabling and a coercive use of bureaucracy was introduced by Adler and Borys (1996) and first applied in management accounting research by Ahrens and Chapman (2004). *Coercive* formalization reflects the typical top-down approach, accentuating centralization and leaving employees with a limited scope of action. *Enabling* formalization, in contrast, *"[…] designs organizational rules that reckon with the intelligence of workers so that formal procedures need not be designed to make the work process foolproof […]"* (Ahrens & Chapman, 2004, p. 279).

Adler and Borys (1996) identify four characteristics that foster an enabling approach to management control: repair, internal transparency, global transparency and flexibility. *Repair* should ensure that users are able to reconfigure processes and react if unforeseen problems occur. With respect to transfer pricing, responsibility centers need to be able to initiate revisions of the transfer pricing system if fundamental problems occur. However, whoever uses the system needs to deeply understand how the system works.

Therefore, the second design principle for an enabling use of bureaucracy is *internal transparency*, which guarantees that users truly comprehend the nature of the system. Only if employees understand the rationale behind a rule can a deeper understanding be achieved. Such an understanding requires a thorough documentation of the transfer pricing system and access to all information related to the transfer pricing system.

In addition to internal transparency, *global transparency* is also necessary to foster an enabling use of bureaucracy. Global transparency requires a broader understanding of a firm's actions and strategies. Users must understand the achieved internal and external objectives of transfer pricing and the effect of their actions not only on their responsibility center but also on the company as a whole. It is argued that such an overall perspective can prevent division managers from taking actions that could harm the company. In an enabling approach, operators have access to extensive information concerning the entire value chain. Therefore, the transfer pricing system helps to clarify the value-creation process of the firm and supports optimizing efforts by division managers to improve the profits not only of their responsibility centers but also of the whole firm.

The last design principle, *flexibility*, indicates that a certain degree of flexibility is needed in a dynamic environment with developing markets and legal amendments. In contrast to repair, which refers to fundamental adaptions and revisions of the transfer pricing system, flexibility in our design refers to exceptions from internal transfer pricing guidelines in particular cases. However, due to the statutory character of international transfer prices, such exceptions need to be carefully justified and documented.

The Adler and Borys (1996) framework has been applied to various settings in management accounting research, such as the study of inter-organizational alliances (Free, 2007), performance measurement systems (Wouters & Roijmans, 2011; Wouters & Wilderom, 2008), transfer pricing and the use of management control systems (Cools et al., 2008) and information-system integration (Chapman & Kihn, 2009). The design of our study is most closely related to Chapman and Kihn's (2009) investigation of the relationships between

information-system integration, enabling control and performance and Cools et al.'s (2008) study on transfer pricing.

2.2.3. Hypotheses Formulation

We draw on findings from case-based research that reveal a high level of integration of tax compliant international transfer pricing into the management control system of a firm (Cools et al., 2008; Rossing & Rohde, 2010) and argue that this integration is positively associated with the perceived success of the transfer pricing system. Such a positive relationship is opposed to the results from analytical research, which typically recommends a decoupling of transfer prices (Baldenius et al., 2004). We argue that firms face enough leeway in the use of a transfer pricing system to overcome this problem. Based on the Adler and Borys' (1996) framework, we distinguish four characteristics of the transfer pricing system that foster its enabling use: repair, internal transparency, global transparency and flexibility. We hypothesize that the relationship between the integration of the transfer pricing system into the management control system and its success is mediated through an *enabling* use of the transfer pricing system. Such an enabling use is necessary to prevent the failures that could arise if tax compliant transfer prices are rigidly applied for management control objectives. Figure 1 displays the structural model underlying the paper.

Insert Figure 1 about here

Our reasoning that the level of integration of the transfer pricing system into the management control system is directly and indirectly (through the enabling use of the transfer pricing system) linked to the perceived success of the transfer pricing system derives from the traditional role of

transfer prices as facilitating and influencing internal decision-making. Transfer prices affect some of the most important management and coordination processes such as budgeting and performance evaluations of division managers, with effects on decision-making and both division and overall company profit. These elements of the management control system are usually portrayed as instruments for achieving decentralization and coordination in multidivisional and multinational firms (see, for example, Horngren et al., 2015, p. 868; Kaplan & Atkinson, 2014; Zimmermann, 2013, pp. 185-188). By providing the relevant data, transfer prices that are integrated into the management control system enhance the transparency of internal decision-making. Compared to decoupled transfer prices, integrated transfer pricing systems are often preferred "both for simplicity and in order to avoid […] any disputes with the tax authorities" (Baldenius et al., 2004, p. 592). In addition, integrated transfer pricing systems are helpful to signal that transfer prices are driven by internal considerations instead of tax optimization (Ernst & Young, 2001, 2003). Taken together, based on Adler and Borys' (1996) framework and contrary to the predictions of analytical models, we expect that the level of transfer pricing system integration into the management control system is positively associated with the perceived success of the transfer pricing system, as formally stated in the first hypothesis:

H1: There is a positive relationship between the level of integration of the transfer pricing system into the management control system and the success of the transfer pricing system as perceived by the corporate-level authorities responsible for transfer pricing.

Drawing on Adler and Borys (1996), we argue that this relationship is partly mediated by an enabling use of the transfer pricing system, which is fostered by the four design characteristics: repair, internal and global transparency, and flexibility. If tax compliant transfer prices are

integrated into the management control system, inconsistencies and failures in the transfer pricing system would be detected and repaired through adjustments of the transfer pricing system. For instance, a cost plus transfer price that does not provide the supplying division with an appropriate profit might reflect inefficiencies in the supplying division or the need for adjustments of the transfer price. Either way, managers would become aware of problematic transfer pricing through the use of transfer prices in the management control system. A similar example is provided by Cools et al. (2008, p. 622). The level of integration is therefore positively associated with the ability to repair the transfer pricing system.

Moreover, if the management control system and thus the decision-making of divisional managers are affected by transfer prices, these managers need to understand how transfer prices are determined, which in turn requires thorough internal documentation of the transfer pricing system and access to the relevant data. Therefore, the level of integration of the transfer pricing system into the management control system is positively associated with internal transparency. In addition, if transfer prices are relevant for management control purposes, managers need to understand not only how the transfer prices are determined but also their role in the broader context, i.e., the overall company perspective. Such a comprehensive understanding is fostered through the global transparency of the transfer pricing system, and we therefore expect to find a positive relationship between the level of integration and global transparency.

The last enabling characteristic relates to the flexible use of the transfer pricing system. In some cases, determining tax compliant and control-relevant transfer prices according to pre-set rules and internal policies might not be feasible. Flexibility refers to the ability to disregard these internal policies and flexibly handle the determination of transfer prices in particular cases. Cools et al. (2008, p. 622) provide an example of such a flexible adjustment of transfer prices

due to market pressure. Other exceptions may relate to the pricing of extraordinarily large orders or the entry into new markets. Taken together, we argue that the level of integration is positively associated with the four design characteristics of repair, internal transparency, global transparency, and flexibility. These relationships are formally hypothesized in H2a-H2d:

H2: There is a positive relationship between the level of integration of the transfer pricing system into the management control system and the enabling use of the transfer pricing system as reflected by the characteristics of (a) repair, (b) internal transparency, (c) global transparency, and (d) flexibility.

Finally, we argue that the enabling use of the transfer pricing system is positively associated with the transfer pricing system's success. If fundamental problems in transfer pricing systems cannot be corrected, the transfer prices become useless for management control purposes. Therefore, we expect repair to be positively associated with transfer pricing system success. Similarly, internal and global transparency is positively associated with transfer pricing system success because these characteristics directly support managers in their decision-making function. Moreover, global transparency ensures that the implications of local processes are visible across the entire company, and thus it fosters the alignment of divisional and corporate goals. Finally, a flexible use of transfer pricing systems is necessary to account for exceptions and thus to prevent deteriorated divisional stewardship. Taken together, the presence of each of the four characteristics is expected to boost the perceived success of the transfer pricing system.

H3: There is a positive relationship between an enabling use of the transfer pricing system as reflected by the characteristics of (a) repair, (b) internal transparency, (c) global transparency,

and (d) flexibility and the transfer pricing system's success as perceived by the corporate-level authorities responsible for transfer pricing.

2.2.4. Research Design and Data

2.2.4.1. Sample Selection and Survey Design

We conducted a questionnaire-based survey according to the basic elements of the tailored design method by Dillman, Smyth, and Christian (2014). During preparation of the survey, a pre-test with managerial accounting, transfer pricing and tax experts was conducted and some questions were adjusted based on the experts' suggestions. The survey was sent to all companies listed on the Swiss Stock Exchange (SIX) in July 2012, except for companies in the financial services sector.[2] All of the 158 companies were first contacted by telephone. After a company agreed to participate in the survey, a questionnaire was sent to an employee in a corporate function with detailed knowledge of the company's transfer pricing system. We focused on employees in corporate functions, as we are interested in the overall approach to transfer pricing. In total, we received 38 completed questionnaires, equaling a response rate of 24 percent. Table 1 presents an overview of our final sample of 38 companies from 9 industry groups.

Insert Table 1 about here

---

[2] Companies from the financial services sector were excluded because of fundamental differences in the value creation process and thus in internal transactions compared with companies from other industries.

The average company has 4,718 employees, ranging between 235 and 26,000 employees (median 2,084). Net sales average 2.033 billion CHF with a range of 40 million CHF to 13 billion CHF. The average international portion of net sales equals 65 percent, emphasizing the importance of transfer prices with respect to profit allocation among business units.

2.2.4.2. Variables Measurement

For the measurement of the four variables reflecting an enabling use of the transfer pricing system as well as of the transfer pricing system's success, we transfer Chapman and Kihn's (2009) measurement approaches into a transfer pricing context. For the measurement of transfer pricing system integration, we rely on literature descriptions of the use of transfer prices for management control purposes (Horngren et al., 2015; Kaplan & Atkinson, 2014; Zimmermann, 2013). All variables are designed based on the results of the principal component analysis (PCA)[3] with oblique Oblimin rotations[4] and Cronbach's alpha as a measure of internal consistency and reliability (Cronbach, 1951). To reflect the theoretical constructs, three independent PCAs are conducted: first, a PCA including the items measuring transfer pricing system integration; second, a PCA addressing the four mediating design characteristics of flexibility, global transparency, internal transparency, and repair; and third, a PCA checking the factorability of items measuring transfer pricing system success. Items that do not match with the theoretical construct are dropped to obtain latent variables that reflect the underlying theoretical model.

---

[3] The Bartlett test of sphericity supports the fit between the model and the correlation matrix for all variables ($p < 0.000$). The Kaiser-Meyer-Olkin measure suggests sampling adequacy for all factor analyses with a measure of sampling adequacy (MSA) value above 0.588.

[4] Because the underlying theoretical model does not assume that the constructs are uncorrelated, the oblique rotation method is the most suitable (Hair, Black, Babin, & Anderson, 2010).

Insert Table 2 about here

Table 2 displays the factor loadings from the three PCAs. All of the factor loadings greater than 0.45 are highlighted in gray. According to Hair et al. (2010, p. 117), factor loadings above 0.5 are considered practically significant, and loadings above 0.70 are described as the "*goal of any factor analysis.*" Except for one item (no. 3), all of the factor loadings are above 0.70, and 16 (out of 19 items) have factor loadings of 0.75 or higher.[5] In addition, the cross loadings are rather small, with the highest (positive) cross loading equaling only 0.321 (item no. 11).

Insert Table 3 about here

For each of the variables, Table 3 presents the general definitions and measurement items used in the questionnaire along with examples from interview data. Based on the results from the PCA (Table 2), the variables are calculated as the mean values of the items that are highlighted in gray in Table 3.

Transfer pricing system integration

We include three of the initial five items to measure the level of integration between the transfer pricing system and the management control system (*TPS integration*). Primarily, a two-component solution was obtained. Because the first component, which includes the majority of

[5] For a sample size of 50, a 0.05 significance level and an 80 percent power level, factor loadings above 0.75 are considered significant (Hair et al., 2010, p. 117). Therefore, the critical value of a factor loading to meet the criteria mentioned would be slightly increased for a sample size of 38. However, many of the factor loadings would still be above or at least be approximately at the threshold value.

the items (including item no. 1, which directly asks for the transfer pricing system integration) explains a higher proportion of the variance (0.51) and obtains a higher Cronbach's alpha (0.51), the second component is dropped. The Cronbach's alpha is rather low, which likely reflects the heterogeneous nature of the variable.[6] The eigenvalue of the factor *TPS integration* equals 1.54. *TPS integration* includes an overall judgment of the level of integration of the transfer pricing system into the management control system, the degree of integration of transfer prices into budgetary planning and the degree of integration of information from cost accounting into the determination of transfer prices.

Characteristics of an enabling use of the transfer pricing system

Repair is measured by including two of the initial four items in the PCA. These items include the ability of those responsibility centers affected by transfer pricing to initiate a revision of the transfer pricing in case of fundamental problems and their ability to hold open discussions on problems with transfer pricing. Because both items load on the same factor, that factor is considered to be repair in the subsequent analysis. Repair has an eigenvalue of 1.40, explains 13 percent of the variance and has a Cronbach's alpha of 0.66.

Measuring *internal transparency* by including all four items appears to be appropriate because they all load on the second factor (called internal transparency in the subsequent analysis). Internal transparency means that responsibility centers affected by transfer pricing know and understand the determination process, that they have access to the relevant data and to the

[6] Hair et al. (2010, p. 125) suggest a Cronbach's alpha of 0.60 for exploratory research as the critical value while stressing the inflation of Cronbach's alpha due to an increase in the number of items. Keeping in mind that only 3 items are used and that a broader range of the sample can further inflate the alpha statistic (Cortina, 1993), an alpha statistic of 0.51 is judged to be low but acceptable.

documentation of the transfer pricing system and that the company keeps detailed records on the transfer pricing system. The eigenvalue of internal transparency equals 3.53, it explains 32 percent of the variance and its Cronbach's alpha is 0.84.

Global transparency is best reflected by two of the initial four items of the questionnaire. These items refer first to the suitability of transfer pricing systems for clarifying the value creation of each responsibility center and second to the potential for each responsibility center to compare its financial performance to that of the others. Global transparency has an eigenvalue of 1.12, explains an additional 10 percent of the variance and has a Cronbach's alpha of 0.62.

We include three of the initial four items to measure the degree of *flexibility* in the transfer pricing system. These items relate to the disregard of internal transfer pricing guidelines and the flexible handling of transfer prices in particular cases. Moreover, we include a reverse-coded item for which no exceptions to the internal guidelines or regulations on transfer pricing are allowed. The factor flexibility has an eigenvalue of 2.29, explains an additional 21 percent of the variance and has a Cronbach's alpha of 0.71.

Taken together, the results from the PCA indicate that the underlying theoretical constructs are appropriately reflected by the data. In particular, only five of the initial 16 items were dropped. The remaining eleven items clearly identify the four assumed mediating design characteristics of an enabling use of the MCS, explaining 76 percent of the total variance.

Perceived transfer pricing system success

Perceived transfer pricing system success (*TPS success*) is measured by five of the initial seven items. These items include assessments of the responsibility centers' overall satisfaction with the transfer pricing system, the cost-benefit analysis, the tax compliance of the transfer pricing

system and the fulfillment of internal and external objectives. The TPS success factor explains 63 percent of the variance, has an eigenvalue of 3.13 and a Cronbach's alpha of 0.85. It includes the overall satisfaction with the transfer pricing system, cost-benefit considerations, the degree of compliance of the transfer pricing system with tax regulations and an evaluation of the fulfillment of the internal and external purposes of transfer pricing.

2.2.4.3. Methodological Approach

Our hypotheses on the relationships between transfer pricing system integration, the four characteristics of an enabling use of the transfer pricing system and the success of the transfer pricing system are investigated through structural equation analysis by applying the partial least squares (PLS) technique. In particular, we rely on the software SmartPLS 3.2.0 for our statistical analyses. PLS has recently gained ground in the accounting literature (e.g., Chapman & Kihn, 2009; Hall, 2008, 2011; Hartmann & Slapničar, 2012; for a review of the use of PLS in management accounting research see Nitzl, 2014) because it is particularly useful for analyzing complex relationships when prior theoretical knowledge on the relationships is limited and sample sizes are rather small. Therefore, it appears to be ideal for estimating the association between transfer pricing system integration and transfer pricing system success while accounting for the mediating effect of an enabling use of the transfer pricing system. Bootstrapping using 5,000 samples with replacement is applied. Following Hair, Hult, Ringle, and Sarstedt (2013, p. 252), the number of bootstrap cases is adjusted to the number of observations in our sample.

PLS analysis is comprised of a measurement and a structural model. To ensure the appropriate interpretation of the results from the structural model, the validity and reliability of the latent

variables (i.e., the measurement model) need to be ensured (Hulland, 1999). As reported in Table 2, the factor loadings are above 0.7 for 18 (out of 19) items, and thus internal reliability is supported (Hair, Ringle, & Sarstedt, 2011). In addition, Table 4 displays common validity and reliability measures for the latent variables of the model. Internal consistency reliability is ensured because the composite reliability of each latent variable is higher than 0.7 (Nunnally & Bernstein, 1994). Convergent validity is guaranteed because the average variance extracted (AVE) is above 0.5 for all of the variables (Hair et al., 2010). Discriminant validity is confirmed because the AVE of each latent variable is higher than its shared variance (squared correlation) with any other construct (Fornell & Larcker, 1981).[7] According to Hair et al. (2011), an item's factor loading should additionally be greater than any of its cross loadings; this condition is also fulfilled (Table 2). Taken together, all of the latent variables meet the validity and reliability criteria.[8]

Insert Table 4 about here

Because both the independent and the dependent variables from the model are obtained from the same survey (answered by the same person in the same measurement context), common method variance (CMV) may bias the results (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003, p. 885). Therefore, we followed several procedures recommended in the literature to avoid CMV as far as possible. *Ex ante*, we tried to reduce CMV by guaranteeing respondents anonymity, using

[7] Table 4 therefore displays the maximum shared variance for each variable.
[8] To check for multicollinearity of the latent variables, the variance inflation factors are calculated. The factors range from 1.10 to 1.44 (with a mean of 1.31) and thus do not indicate any problems (threshold values are between 5 and 10).

neutral wording, reducing comprehension problems by wording questions as precisely as possible[9] (Harrison, McLaughlin, & Coalter, 1996; Lindell & Whitney, 2001) and integrating some questions worded with an opposing orientation. *Ex post*, the mediating role of our design characteristics adds complexity to our model, reducing the influencing potential of CMV. Moreover, we perform some statistical tests in order to ensure that our findings are not solely driven by CMV.

First, we employ Harman's single-factor test by performing a PCA (without rotation) including all items of our model.[10] Following Podsakoff and Organ (1986), a single resulting factor or a factor accounting for the majority of covariance between the items would signal the prevalence of CMV. However, the PCA reveals six factors with eigenvalues greater than one, and the first factor accounts for 32% of total variance only, indicating that CMV is not pervasively affecting our results. Second, another promising approach to control for CMV is the use of a marker variable (Lindell & Brandt, 2000; Lindell & Whitney, 2001). A marker variable should theoretically not be correlated to the other variables (at least to one) used in the model; if there are observed correlations between the marker variable and the other variables, these are assumed to be caused by CMV. Such a procedure allows researchers to parcel out the effect of CMV from observed relationships (Podsakoff et al., 2003). Therefore, using financial performance as a marker variable,[11] we test the biasing effect of CMV on our results. After controlling for CMV, the correlations between TPS integration and TPS success remain in a range between

[9] Therefore, a pretest of the survey was conducted with experts in the topic in order to adjust any questions that may have initially been difficult to understand according to their suggestions.
[10] Note that the factor analysis includes all 19 items used to measure TPS integration, repair, internal transparency, global transparency, flexibility and TPS success.
[11] Financial Performance is measured by three items and has a Cronbach's alpha of 0.95. The items ask for the respondents' assessments of the firm's return on sales, sales growth and financial performance in comparison to the firm's competitors.

0.37 and 0.49, the correlations between repair and TPS success remain in a range between 0.52 and 0.62, and the correlations between internal transparency and TPS success remain in a range between 0.54 and 0.62.[12] Considering these stable and highly significant values, we conclude that CMV does not primarily drive our results.

2.2.5. Results

2.2.5.1. Descriptive Statistics

The descriptive statistics for our model variables are presented in Table 5. Panel A of Table 5 gives an overview of the raw data. Variables are calculated as the mean values of the items and theoretically range between 1 and 7. On average, the transfer pricing system appears to be well integrated, and firms perceive their transfer pricing systems as being rather successful (mean values of 5.00 and 5.08, respectively). This finding supports our reasoning that transfer prices are indeed (well) integrated into the management control system and not decoupled. Note that we only ask for *international* transfer prices and that the sample companies have, on average, 65 percent of total sales internationally. In contrast to these rather high mean values for TPS integration and TPS success, the mean values of the four design characteristics are lower, ranging between 3.84 and 4.76. Within these four variables, we observe the highest mean value for repair, indicating that transfer pricing systems are generally repairable. The lowest mean value corresponds to global transparency, indicating that the transparency of transfer prices and of their effects across different responsibility centers is rather limited. Moreover, the flexible use of the transfer pricing system is not widespread, which may reflect concerns regarding potential disputes with the tax authorities.

[12] Note that there are different procedures to identify the concrete value of CMV for a correlation, which in turn result in different values for the respective correlation coefficients.

Insert Table 5 about here

For further analyses, variables are computed as the average standardized response of all items loading above 0.45 on a particular factor (Table 5, Panel B), which is considered to be an appropriate critical value for exploratory research (Tabachnick & Fidell, 2001).[13]

Panel C of Table 5 presents Pearson correlation coefficients for these standardized variables. As expected, TPS integration is positively and significantly correlated with internal transparency, repair and TPS success. Moreover, TPS success is also positively and significantly correlated with repair and internal transparency. While there are also significant relationships between our four design variables for an enabling use of the *transfer pricing system* (in particular between repair and internal transparency, global transparency and flexibility as well as between global transparency and repair), the corresponding variance inflation factors do not indicate multicollinearity problems. Overall, the correlation statistics provide univariate evidence that supports our expected relationships between TPS integration, an enabling use of the transfer pricing system and TPS success. In contrast to our expectations, global transparency is negatively correlated with TPS success, but this relationship is not significant.

2.2.5.2. Results from Structural Equation Modeling

Our structural equation model tests the hypotheses that transfer pricing system integration is positively related to transfer pricing system success (H1), that transfer pricing system integration is positively related to an enabling use of the system as reflected by repair, internal transparency, global transparency, and flexibility (H2a-H2d) and that an enabling use of the

[13] Even lowering this critical value to 0.35 does not impact our results with respect to the variable compositions.

transfer pricing system as reflected by its four design characteristics is positively related to its success (H3a-H3d). Table 6 presents our results on path coefficients, corresponding levels of statistical significance, multiple squared correlations (mult. R^2) and effect sizes (f^2).

Insert Table 6 about here

With respect to the first hypothesis H1, the results of the PLS analysis indicate a positive association between the integration of the transfer pricing system into the management control system and transfer pricing system success (Table 6, Panel A). However, the link is significant at the 0.10 level only (one-tailed test). Therefore, we find only weak support for hypothesis 1. Nevertheless, the total effect of TPS integration on TPS success – including the indirect (mediating) effects through internal transparency, global transparency, repair, and flexibility – is substantial and highly significant, thereby emphasizing the overall importance of transfer pricing system integration (coefficient of 0.478, p-value of 0.000, untabulated). Thus, as assumed by our research design, the association between TPS integration and TPS success appears to be mediated through internal transparency, global transparency, repair, and flexibility.

Regarding the second set of hypotheses on the relationship between TPS integration and the four characteristics of an enabling use of the transfer pricing system (H2a-H2d), the results of the PLS analysis indicate positive and significant associations for three of the four design variables (Table 6, Panel A). In particular, TPS integration is positively related to repair (0.336, $p<0.05$), internal transparency (0.388, $p<0.01$), and global transparency (0.246, $p<0.10$). These findings strongly support hypotheses H2a, H2b and weakly support H2c. While the relationship between

TPS integration and flexibility is also positive, it is not significant and thus does not support hypothesis H2d (0.183, $p>0.10$).

Panel B of Table 6 presents the results for the third set of hypotheses (H3a-H3d) along with the effect sizes and the multiple squared correlations for TPS success. Consistent with H3a and H3b, there are significant positive relationships between repair and TPS success (0.450, $p<0.01$) as well as between internal transparency and TPS success (0.321, $p<0.05$). However, global transparency and flexibility both have a negative (but not significant) association with TPS success. Therefore, we find only partial support for our third set of hypotheses. The multiple R^2 for TPS success equals 0.605, indicating that TPS integration and the four variables of an enabling use explain 61 percent of the variance of TPS success (adjusted $R^2 = 0.543$). Compared with similar studies (e.g., Chapman & Kihn, 2009; Hall, 2008, 2011), this value can be considered rather high, thereby indicating a good fit for our model.

Overall, the results of the PLS analysis partially support our theoretical model. In particular, our results provide evidence not only for a direct positive relationship between TPS integration and TPS success but especially for indirect positive relationships through the characteristics repair (0.163, $p=0.065$) and internal transparency (0.129, $p=0.076$). This result is also reflected by a Cohen's effect size f^2, equaling 0.133 for TPS integration, 0.360 for repair and 0.181 for internal transparency.[14] In other words, stronger integration of the transfer pricing system into the management control system is associated with the higher success of the transfer pricing system, and this relationship is further enhanced through the enabling characteristics repair and internal transparency. However, although statistically not significant, we find negative associations

[14] Note that effect sizes of 0.02, 0.15 and 0.35 are interpreted as small, medium and large according to Cohen (1988).

between global transparency (as well as flexibility) and TPS success. One potential reason could be that the increased transparency of the transfer pricing system across different responsibility centers generates additional conflicts between the managers of these responsibility centers because the value creation of each responsibility center becomes more visible. The insignificant result for flexibility might reflect concerns about potential disputes with the tax authorities that are associated with a flexible use of the transfer pricing system.[15]

2.2.5.3. Robustness Analyses

We perform several additional analyses to investigate the robustness of our survey-based findings. First, we address concerns regarding the computation of our variables and our approach of separately calculating the measurement and the structural model. There is a controversy in the literature regarding the different methods for estimating factor scores. In our analysis, we rely on the common approach of computing the variables as the average standardized response of items (see also Chapman & Kihn, 2009). Alternatively, we use regression-based factor scores as an initial sensitivity check for our results. Following such a procedure means that a factor is no longer described as the average of the standardized items with high loadings only but is rather described by each item, while the weight of an item directly depends on the size of its respective factor loading (Hair et al., 2010, p. 117). The results (untabulated) of the path model analysis based on bootstrapping with 5,000 drawings using regression-based factor scores are similar to our initial results. In particular, there are no differences with respect to the direction, but there are differences with respect to the magnitude for some of the estimated effects, confirming our main findings.

[15] For the negative (and not significant) associations between global transparency (as well as flexibility) and TPS success, see also our affirmative testimonies from interview data below.

Second, we conduct a PLS analysis using the items from the main model before standardization. With this procedure, SmartPLS optimizes the complete (i.e., measurement and structural) model including the weighting of the single items. Again, the results (untabulated) are similar to our main analyses with one exception. The direct relationship between TPS integration and TPS success becomes insignificant, and TPS integration is thus positively associated with TPS success only through an enabling use of the transfer pricing system.

Third, our initial step-wise procedure of first performing principal component analyses and second testing our hypotheses based on the PLS technique is criticized by some researchers (Nitzl & Chin, 2016). Therefore, we next optimize the complete model in SmartPLS including the selection of the unstandardized items used to specify the latent variables of our model. The results of this analysis are presented in Table 7. Following such a procedure slightly changes the measurement of global transparency and flexibility, but the main results remain robust. In particular, TPS success is still positively associated with repair and internal transparency, while the total effect of TPS integration on TPS success – including the indirect (mediating) effects – remains unchanged (coefficient of 0.478, p-value of 0.000, untabulated). Taken together, results from these robustness analyses indicate that neither the computation of our variables nor the step-wise procedure substantially impacts our main findings.

Insert Table 7 about here

Next, one could argue that our results may be biased due to omitted variables, and we therefore include two additional control variables that may impact TPS success, namely the number of

profit centers (#Profit centers) and the proportion of foreign sales (Foreign sales).[16] The number of profit centers is a rough proxy for the number of internal transactions, and the proportion of foreign sales is a rough proxy for the degree of internationalization of a company. Taken together, these variables are a crude indicator of the importance of internal cross-border transactions for the whole company. Results from this additional analysis are presented in Table 8. While there are no substantial differences in the associations between TPS integration, the characteristics of an enabling use and TPS success, both control variables are positively associated with TPS success. One potential reason for this finding could be that the higher importance of internal cross-border transactions translates into more sophisticated transfer pricing systems, which in turn are associated with higher TPS success.

Insert Table 8 about here

Next, we re-run our analyses based on an alternative measurement of our dependent variable *TPS success* that captures the perceived conflicts between various transfer pricing objectives. Our new variable, *TPS conflicts,* is based on five of seven initial items that measure the level of perceived conflict between transfer pricing objectives. Primarily, a two-component solution was obtained. Again, we use the factor that includes the majority of items (including the item that directly asks for the overall level of conflict between transfer pricing objects). The factor *TPS conflicts* explains 51 percent of the variance and has an eigenvalue of 2.57 and a Cronbach's alpha of 0.76. In addition to an overall judgment of conflicts between different transfer pricing objectives, it further includes particular assessments of conflicts between coordination and

[16] Note that we include the standardized values of these variables in the PLS analysis.

internal profit determination, coordination and tax optimization, internal profit determination and tax compliance as well as internal profit determination and tax optimization. The results from the PLS analysis based on bootstrapping with 5,000 drawings are displayed in Panel A and Panel B of Table 9.

Insert Table 9 about here

Consistent with our expectations, the path coefficient from TPS integration to TPS conflicts is negative but not significant. However, the total effect, including the indirect (mediating) effects, becomes significant at the 10 percent level (coefficient of -.255, p-value of 0.065, untabulated). The results thus suggest that the integration of the transfer pricing system into the management control system is associated with reduced conflict between different transfer pricing objectives. Again, similar to our baseline model, the relationship is mediated through the design variables of the transfer pricing system. However, the relationship between TPS integration and TPS conflicts appears to be much weaker than the relationship between TPS integration and TPS success. One reason for this observation might be that we dropped variables related to internal profit determination when considering the measurement of TPS integration but included them when considering TPS goal conflicts. This could also be an explanation for the much lower multiple R^2 (0.212) in this new model. Panel B of Table 9 presents the results for the third set of hypotheses (H3a-H3d). Surprisingly, the path coefficient between repair and TPS conflicts becomes positive. However, the effect is not significant (coefficient of 0.160, p-value of 0.244). In contrast, the results suggest that the path between internal transparency and TPS conflicts, both in terms of magnitude and significance, is the strongest. This finding is also reflected by

obtaining the highest Cohen's effect size f^2 for internal transparency. Therefore, a higher degree of internal transparency appears to be associated with a lower level of conflict between transfer pricing objectives. By contrast, global transparency and flexibility appear to be unrelated to TPS goal conflicts. Overall, the results are generally consistent with our findings from the main model (except for repair).

Finally, we estimate our model by ordinary least square (OLS) analysis because most researchers are more familiar with this technique when it comes to the estimation of relationships. To test the structural model of our hypotheses, we run six independent OLS analyses. The results are presented in Table 10. Although the estimated coefficients and the t-statistics differ from our PLS analysis (due to the different estimation model), the results are similar to the results from our main model. In particular, we obtain positive and significant associations between TPS integration, repair and internal transparency and TPS success. Not controlling for the use of the transfer pricing system, the total effect of TPS integration on TPS success equals 0.530 and is highly significant (p-value of 0.002). However, only 55 percent (0.291/0.530) of this total effect can be explained by the direct effect of TPS integration, while 45 percent of the total effect can be attributed to the mediating variables. As in our PLS model, the indirect effects of TPS integration on TPS success through repair (coefficient of 0.168, p-value of 0.037) and internal transparency (coefficient of 0.138, p-value of 0.043) remain similarly substantial and highly significant.

Insert Table 10 about here

Taken together, the results from these additional analyses unanimously support the robustness of our main results.

2.2.5.4. Affirmative Testimonies from Interview Data

To further substantiate our findings, we provide additional insights into the transfer pricing systems of three multinational enterprises. In particular, we illuminate potential conflicts and problems in the use of transfer pricing and the perceived success of the transfer pricing system.

We use theoretical sampling (Eisenhardt, 1989) to identify useful cases for our investigation based on the following criteria. First, we limit our potential sample to the manufacturing industry to enhance the comparability of our findings. Second, we focus on multinational enterprises with a substantial amount of internal cross-border transactions, as only these enterprises theoretically face a goal conflict between tax compliance and management control. Third, we aim to use enterprises that have implemented a single set of transfer pricing books. Again, only in this case does the goal conflict theoretically become evident. Moreover, focusing on one set of books enhances the comparability of our findings.

Within each enterprise, we conducted one interview that lasted, on average, two hours. Regarding the interview subjects, we focused on those persons responsible for the transfer pricing system at the corporate level. In addition, we interviewed a divisional manager (company A) and a corporate controller (company C). The interviews were conducted between October and December 2016. The semi-structured interviews were guided by questions based on our survey questionnaire and our results. The questions were open enough to allow for unexpected findings. For reasons of simplicity, we focused on internal transactions involving tangible assets and excluded the licensing of intellectual property (IP) from our analysis. The interviews were tape-recorded and transcribed.

Panel A of Table 11 provides an overview of the sample companies A, B, and C with respect to the firm size in terms of sales, internationalization, organizational structure, and value chain.

Insert Table 11 about here

Panel B of Table 11 outlines the transfer pricing systems of the three case companies. The organizational structure of company A with its various profit centers is reflected in a decentralized approach to transfer pricing. The central guideline on transfer pricing only states that transfer prices must comply with the arm's length principle. Transfer prices are negotiated between profit centers (based on price lists or individual prices for customer specific products). This approach of internal negotiations on the transfer price reflects the negotiating power and the functions and risks of the involved profit centers. The transfer pricing systems of companies B and C are rather contrary to the decentralized approach to transfer pricing seen in company A. Both companies have comprehensive and detailed transfer pricing policies at the corporate level that govern the determination of transfer prices. These guidelines or policies determine the transfer pricing method with respect to typical intra-company transactions including the cost basis, markups, margins, and handling of currency conversions.

Next, we provide descriptive evidence for the (enabling) use of a (single) transfer pricing system, which must simultaneously fulfill tax compliance and management control purposes. Based on the findings from analytical research, which recommends the decoupling of transfer prices to avoid conflicts between tax compliance and management control, we directly asked our interviewees about the existence of such goal conflicts in their companies. Company A's interviewees observe no direct goal conflict between the two objectives; on the contrary,

company A is convinced that tax authorities accept the negotiated transfer prices as being tax compliant because they correspond with their decentralized approach to controlling the group. Company A's interviewee stated: "*With respect to tax compliance, the total result of the legal entity is crucial. If this result is feasible, there are usually no further discussions with the tax authorities. […] Moreover, keeping one set of books is a benefit with respect to tax compliance, since we can always argue that there are not manipulations for tax optimization purposes*." However, the decentralized approach of negotiating transfer prices is accompanied by the high potential for conflict between the negotiating parties. If the internal transactions occur within the same business unit, the internal conflicts are solved by the business unit manager. If different business units are involved, potential conflicts are solved by division or corporate representatives. Moreover, there is an arbitration panel that resolves conflicts that occur between the involved parties *after* transfer prices are determined.

Company B's interviewee also observes no conflict between tax compliance and management control: "*Because the transfer pricing system has been designed in accordance with basic economic principles and to facilitate decentralized decision-making, it is in accordance with tax regulations. … Thus, we assume that if our business decisions are reasonable, then these decisions must also be reasonable with respect to tax compliance.*" The high integration of the transfer pricing system with the management control system in combination with the ability to repair the transfer pricing system facilitates internal decision-making and reduces potential conflicts between the involved parties. For instance, transfer prices are integrated into the budgeting process of the company. If the planning results in a loss for some legal entities due to the planned transfer prices, the problem is escalated to a transfer pricing committee, which can revise the transfer pricing system accordingly. Due to the integration with the budgetary

planning, such problems can be solved in advance. The company also provides insights into the practical handling of another problem of transfer pricing: the obfuscation or "swelling" of the internal cost structure that results from the use of the cost plus method in multi-stage production processes. Such obfuscation is prevented by applying the markup of the transfer price solely to the production costs of each production center instead of to the total costs.[17]

Last, but not least, company C's interviewees perceive no conflict between management control and tax compliance for two reasons. First, the company's primary objective with respect to transfer pricing is tax compliance. Second, in addition to tax compliance, the main objective of the transfer pricing system is "*to not distort decentralized decision-making.*" To ensure this, transfer prices do not affect the performance evaluation of functional center managers because they are evaluated based on the business unit's residual income. The transfer pricing guideline is detailed and comprehensive. The low degree of flexibility results in few conflicts between functional centers, business units or legal entities. Within each business unit, there are transfer pricing officers who are responsible for adherence to the transfer pricing guideline. Problems are discussed with the corporate tax department on a monthly basis. In the case of severe conflicts, there is an escalation to the management board of the company, which ultimately resolves the conflict.

To gather more insights into the finding that neither global transparency nor flexibility is found to significantly mediate the relationship between transfer pricing system integration and transfer pricing system success, we directly asked our interviewees about these relationships in their enterprises. The results from our interviews reveal differences in the role of flexibility and

[17] The total costs include material and production costs. Material costs in turn include the costs of internal products (i.e., the transfer prices paid to the upstream production centers).

global transparency between centralized and decentralized transfer pricing systems. As expected, the interviewees from the companies with a centralized transfer pricing system emphasize that flexibility is detrimental to tax compliance and thus to the perceived success of the transfer pricing system. In contrast, in the company taking a decentralized approach, flexibility is, by definition, part of the transfer pricing system and thus important for its success. However, it must be noted that the central transfer pricing policy of the decentralized company only states that transfer prices must comply with the arm's length principle, which allows for much more flexibility than a comprehensive and detailed internal transfer pricing policy. Another discrepancy is that, for the two companies with a more centralized approach, both internal and global transparency are important for the success of their transfer pricing system. For the company with the decentralized approach, only internal (but not global) transparency is perceived as being important. Thus, the different level of decentralization seen in the companies in our sample might explain the insignificant results for global transparency and flexibility that we obtain from the PLS analysis. In none of the case companies did we find severe conflicts between management control and tax compliance, which is partly due to the high integration between the transfer pricing system and the management control system.

In sum, there is substantial descriptive evidence supporting the presumed enabling use of a transfer pricing system through which firms are able to fulfill tax compliance and management control at the same time. However, the short rudimentary descriptions illustrate the complexity of transfer pricing systems in practice and of their integration into the management control system. All of the interviewees agreed that monitoring and documenting transfer prices on a transaction-by-transaction basis is impossible due to the huge number of transactions. Tax audits therefore examine the tax compliance of the firms' transfer pricing systems from an overall

perspective and examine the appropriateness of the global profit allocation and of the internal transfer pricing policy.

2.2.6. Conclusions

This paper examines how the integration of a single-book tax compliant transfer pricing system into the management control system affects the perceived success of the transfer pricing system. To this end, we focus on the overall transfer pricing system (instead of on transfer pricing methods as applied to single transactions) and ask how the integration of the transfer pricing system into the management control system is associated with the enabling use of the transfer pricing system and how these interrelationships translate into greater perceived success.

While results from the analytical transfer pricing literature suggest the decoupling of transfer prices to overcome conflicts between tax compliance and management control, we draw on the theoretical framework of Adler and Borys (1996) and suggest that potential conflicts may be overcome by an enabling use of the transfer pricing system. In particular, we hypothesize that the level of transfer pricing system integration is positively correlated with its success and that this relationship is mediated through an enabling use of the system, which is reflected in the characteristics of repair, internal transparency, global transparency, and flexibility.

Based on a sample of 38 Swiss multinational enterprises, the results from structural equation modeling using the PLS technique partly support our reasoning. In particular, transfer pricing system integration is indeed positively associated with transfer pricing system success, and this relationship is mediated through repair and internal transparency. Both characteristics support an enabling use of the transfer pricing system because managers can truly understand how the transfer prices are determined and are able to intervene if fundamental problems occur.

Obviously, these characteristics help to avoid deteriorated management control in an integrated transfer pricing system. However, neither global transparency nor flexibility is found to significantly mediate the relationship between transfer pricing system integration and transfer pricing system success. In the case of transfer pricing, it thus appears that global transparency, due to the enhanced comparability of value creation across responsibility centers, potentially fuels conflicts between responsibility centers and thus is negatively related to transfer pricing system success. With respect to flexibility, we believe that the disregard of internal transfer pricing policies may be detrimental to the purpose of tax compliance and is thus negatively associated with transfer pricing system success. Instead firms may aim to generally describe exceptions known ex ante in the transfer pricing policies or to adjust the guidelines to formally account for deviations.

Our survey-based findings are substantiated by additional insights into the transfer pricing systems of three multinational enterprises. The case companies are multinational companies in the manufacturing industry with substantial internal cross-border transactions that use a single set of transfer pricing books. One company applies a decentralized approach of negotiated transfer prices, while the other two companies apply a centralized approach with detailed transfer pricing guidelines that govern the determination of transfer prices. Tax compliance is the primary transfer pricing objective for all three firms. As expected, we find internal transparency and the potential to repair the transfer pricing system in the case of fundamental problems to be important factors in the success of integrated transfer pricing systems. However, the benefit of global transparency and of the flexible use of the transfer pricing system might depend on the degree of centralization in the approach. Whereas greater global transparency and

less flexibility appear to be beneficial for centralized approaches, the opposite may be true for decentralized companies.

Nevertheless, our results should be interpreted cautiously, as our study is subject to several limitations, which give rise to future research possibilities. First, all variables of interest are latent constructs, and our study thus requires the development of new measurements. Although the validity and reliability of our latent variables are generally satisfactory, future research could further improve some of our measurements, in particular, that of transfer pricing system integration. Moreover, both dependent and independent variables are obtained through the same survey, creating the potential for CMV. We investigate this concern in the robustness section of this paper (section 5.3) and find no indication that our results are severely biased due to CMV. In addition, our questionnaire is addressed to employees at the corporate level, which allows us to capture an overall view on transfer pricing. Future research could involve multiple points of view from the same company by addressing both managers at the corporate level and profit center managers.

Second, our results are based on a limited number of observations from one country. However, because national transfer pricing legislation in Switzerland completely adheres to that of the OECD and there is no supplementary transfer pricing legislation, the restriction to one country could also enhance the generalizability of our results across countries without specific national transfer pricing legislation. Nevertheless, further research could investigate whether our findings also hold for companies from other countries.

Third, to the best of our knowledge, this is the first empirical study on transfer pricing that applies to the totality of intercompany transactions instead of focusing on single transactions.

While such an approach enables us to provide a more holistic picture of transfer pricing in practice, we naturally depart from common research ground. Despite this limitation, we are confident that our study adds a new and important perspective to the ongoing discussion on transfer pricing. Both survey-based and case-based research could build on our model and more closely investigate the relationships between transfer pricing system integration and an enabling use of the transfer pricing system. Moreover, the results from our interviews indicate that it might be particularly fruitful to consider the level of centralization of the transfer pricing system when further investigating the integration of the transfer pricing system into the management control system.

Table 1. Sample distribution by industry group

By industry	N
Mechanical and plant engineering	8
Pharma/medical and bio technology	5
Transport/logistics/public transportation	5
Building and raw materials/chemistry/synthetics	5
Consumer goods	4
Utilities	2
Electronics	2
Telecommunications	1
Other	6
Total	38

Table 1 reports the sample distribution by industry group.

Table 2. Principal component analysis (Oblimin rotated)

No.	Item	TPS Integration	Repair	Internal Transparency	Global Transparency	Flexibility	Perceived TPS Success
1	Our transfer pricing system is fully integrated into the management control system of our company.	0.765					
2	Transfer prices are incorporated as cost or sales prices into the budgetary planning of our profit centers.	0.762					
3	We integrate information from cost accounting into the determination of transfer pricing.	0.607					
4	If fundamental problems occur, the responsibility centers (e.g., profit centers) affected by transfer pricing initiate a revision of the transfer pricing.		0.806	0.125	0.070	0.054	
5	We openly discuss problems with transfer pricing in our company.		0.742	0.197	0.205	-0.053	
6	Responsibility centers (e.g., profit centers) affected by transfer pricing know and understand how transfer prices are determined in our company.		0.284	0.799	-0.063	0.032	
7	Responsibility centers (e.g., profit centers) affected by transfer pricing have access to the relevant data (e.g., costs) that affect transfer prices.		-0.031	0.812	-0.082	0.083	
8	Our company keeps detailed records on our transfer pricing system.		0.266	0.750	-0.271	-0.065	
9	Responsibility centers (e.g., profit centers) affected by transfer pricing have unlimited access to documentation on our transfer pricing system.		-0.041	0.800	0.219	-0.041	
10	Our transfer pricing system helps to clarify the value creation of each responsibility center.		0.013	0.124	0.832	0.147	
11	Due to our transfer pricing system, we are able to compare the financial performance of each responsibility center with that of other centers.		0.321	-0.377	0.778	-0.079	
12	In particular cases, we can disregard our internal guidelines on transfer pricing to determine transfer prices.		-0.337	0.130	0.250	0.808	
13	In particular cases (e.g., major orders), our transfer pricing is flexibly handled.		0.251	0.200	0.041	0.768	
14	There are no exceptions to our internal guidelines/regulations on transfer pricing (reverse-coded item).		0.176	-0.448	-0.186	0.791	
15	Overall, we are very satisfied with our transfer pricing system.						0.836
16	Overall, the benefits of our transfer pricing system outweigh the costs.						0.788
17	Our transfer pricing system complies completely with tax regulations.						0.717
18	Overall, our transfer pricing system achieves our internal objectives for transfer pricing completely.						0.835
19	Overall, our transfer pricing system achieves our external objectives for transfer pricing completely.						0.774

Table 3. Variable definitions

Variable	
TPS Integration	General definition: The integration of a tax compliant transfer pricing system into the management control system in terms of the scope and intensity of using tax compliant transfer prices for management control purposes Operationalization for the questionnaire: • Our transfer pricing system is fully integrated into the management control system of our company. • Transfer prices are incorporated as cost or sales prices into the budget planning of our profit centers. • We integrate information from cost accounting into the determination of transfer pricing. • Our transfer prices influence the performance evaluation of our profit centers. • Our transfer prices influence (indirectly) the variable compensation of the management of the responsibility centers (e.g., profit centers). Examples/quotes from interview data: • Tax compliant transfer prices are part of cost accounting, the budget planning of divisions, the profitability measure for the value chain, the EBIT of a legal entity (Firm A) • The tax compliant transfer price is part of the OVC (Operating Value Contribution) (Firm C)
Repair	General definition: Inconsistencies and failures in the transfer pricing system would be detected and repaired through adjustments of the transfer pricing system; for instance, a cost plus transfer price that does not provide the supplying division with an appropriate profit might be adjusted Operationalization for the questionnaire: • Ideas for improvement on transfer pricing by the responsibility centers (e.g., profit centers) affected by transfer pricing are very welcome. • If fundamental problems occur, the responsibility centers (e.g., profit centers) affected by transfer pricing initiate a revision of the transfer pricing determination. • If original conditions change, the responsibility centers (e.g., profit centers) affected by transfer pricing initiate a revision of the transfer pricing determination. • We openly discuss problems with transfer pricing in our company. Examples/quotes from interview data: • If fundamental problems occur, divisions or business units can initiate a revision (Firm A) • Most importantly, the transfer price must enable the sales unit to make sales, more precisely, the sales they should make (Firm A) • If there is a problem (or a change in the business model), then it's discussed in the TPS working groups and/or in the TPS board. Working groups have a meeting once a week or monthly depending on the agenda; the TPS board meets monthly. Decisions may lead to organizational changes or a change of the TPS policy (Firm B) • If, in exceptional cases, well-founded deviations from the TPS policy occur, then the deviation will be discussed in the tax meetings and usually the deviation will become a standard (Firm C)

Variable	
Internal Transparency	General definition: The transfer pricing system is comprehensively documented; (decentralized) management understands how transfer prices are determined and has access to all relevant data Operationalization for the questionnaire: • Responsibility centers (e.g., profit centers) affected by transfer pricing know and understand how transfer prices are determined in our company. • Responsibility centers (e.g., profit centers) affected by transfer pricing have access to the relevant data (e.g., costs) that affect transfer prices. • Our company keeps detailed records on our transfer pricing system. • Responsibility centers (e.g., profit centers) affected by transfer pricing have unlimited access to documentation on our transfer pricing system. Examples/quotes from interview data: • There is transparency in the plants and the sales units (Firm A) • There is transparency in the legal entities (Firm B) • Transfer pricing methods and markups are public information within the group (Firm B)
Global Transparency	*General definition*: (Decentralized) management understand the role of (tax compliant) transfer prices in the broader context, i.e., the overall company perspective *Operationalization for the questionnaire*: • Responsibility centers (e.g., profit centers) affected by transfer pricing know and understand the achieved internal objectives of transfer pricing in our company. • Responsibility centers (e.g., profit centers) affected by transfer pricing know and understand the achieved external objectives of transfer pricing in our company. • Our transfer pricing system helps to clarify the value creation of each responsibility center. • Due to our transfer pricing system, we are able to compare the financial performance of the responsibility centers. *Examples/quotes from interview data:* • There is no global transparency since the ERP systems are not connected. There is no transparency about value chain profitability (Firm A) • There is global transparency across business sectors and business divisions (Firm C)
Flexibility	*General definition*: Flexibility refers to the ability to disregard the internal transfer pricing policies and to flexibly handle the determination of transfer prices in particular cases *Operationalization for the questionnaire*: • Our transfer pricing system strictly complies with internal guidelines on transfer pricing. • In particular cases, we can disregard our internal guidelines on transfer pricing to determine transfer prices. • In particular cases (e.g., major order), our transfer pricing is flexibly handled. • There are no exceptions to our internal guidelines/regulations on transfer pricing. *Examples/quotes from interview data:* • Flexibility is rather low. Only a small minority of internal transactions are allowed to deviate from our transfer pricing policy. Such deviations must be discussed and agreed upon by the corporate tax department (Firm B) • There is almost no possibility to deviate from our transfer pricing policy; besides, there is a strong tendency over the previous years to rule out any deviations from the transfer pricing policy (Firm C)

Variable	
Perceived TPS Success	*General definition*: Perceived transfer pricing system success is the assessment of overall satisfaction with the tax compliant transfer pricing system and, in particular, the satisfaction with how tax compliant transfer prices fulfill management control objectives *Operationalization for the questionnaire*: • Overall, we are very satisfied with our transfer pricing system. • Overall, the benefits of our transfer pricing system outweigh the costs. • Our transfer pricing system complies completely with tax regulations. • Overall, our transfer pricing system achieves completely our internal objectives for transfer pricing. • Overall, our transfer pricing system achieves completely our external objectives for transfer pricing. *Examples/quotes from interview data:* • Success is defined by tax compliance, good decision-making and few conflicts (Firm A, B and C) • We are fairly pleased with our transfer pricing system since it works: there are only few conflicts with tax authorities and at the same time, the transfer pricing system doesn't disturb management control and business decision-making (Firm C)

Table 3 provides (general) definitions, an operationalization (used in the questionnaire) and practical examples of all variables. Those items used in the second step of our baseline model (initial step-wise procedure of first performing PCA and second, testing our hypotheses based on the PLS technique) are highlighted in gray.

Table 4. Validity and reliability measures of variables used in the PLS analysis

Variable	Composite reliability	Average variance extracted	Maximum shared variance
TPS Integration	0.752	0.504	0.229
Repair	0.855	0.746	0.368
Internal Transparency	0.895	0.682	0.389
Global Transparency	0.833	0.716	0.086
Flexibility	0.724	0.510	0.075
Perceived TPS Success	0.893	0.626	0.389

Table 4 reports validity and reliability measures for all of the variables used in the PLS analysis. The first column shows the composite reliability, and the second column shows the average variance extracted of a variable. Column three shows each variable's highest shared variance with any of the other variables.

Table 5. Descriptive statistics

Panel A: Descriptive statistics for the untransformed variables

Variable	Observations	Mean	Std Dev	Min	Max
TPS Integration	38	5.00	1.31	2.00	7.00
Repair	38	4.76	1.38	2.50	7.00
Internal Transparency	38	4.53	1.49	1.75	6.75
Global Transparency	38	3.84	1.42	1.00	7.00
Flexibility	38	4.18	1.63	1.00	7.00
Perceived TPS Success	38	5.08	1.12	2.00	6.80

Panel B: Descriptive statistics for the standardized variables

Variable	Observations	Mean	Std Dev	Min	Max
TPS Integration	38	0.00	0.71	-1.71	1.08
Repair	38	0.00	0.84	-2.08	1.59
Internal Transparency	38	0.00	0.83	-1.62	1.23
Global Transparency	38	0.00	0.85	-1.72	1.90
Flexibility	38	0.00	0.80	-1.57	1.39
Perceived TPS Success	38	0.00	0.79	-2.19	1.23

Panel C: Pearson correlation coefficients

	1)	2)	3)	4)	5)	6)
1) TPS Integration	1					
2) Internal Transparency	0.388**	1				
3) Flexibility	0.183	0.061	1			
4) Repair	0.336**	0.426***	0.172	1		
5) Global Transparency	0.246	-0.054	0.274*	0.293*	1	
6) Perceived TPS Success	0.478***	0.624***	0.033	0.607***	-0.039	1

Table 5 presents descriptive statistics for the untransformed (Panel A) and the standardized variables (Panel B) as well as correlation statistics (Panel C). All statistics are presented for the full sample of 38 firms. Panel C reports bivariate Pearson correlation coefficients. ***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent levels for a two-tailed test of statistical significance, respectively.

Table 6. Results of the PLS analysis

Panel A: Results for hypotheses H1 and H2a-H2d

Paths to	Paths from				
	TPS Integration	std. error	t-value	p-value	mult. R^2
Repair	0.336**	0.178	1.895	0.029	0.113
Internal Transparency	0.388***	0.157	2.465	0.007	0.151
Global Transparency	0.246*	0.175	1.408	0.080	0.061
Flexibility	0.183	0.171	1.068	0.143	0.033
Perceived TPS Success	0.263*	0.170	1.547	0.061	

Panel B: Results for hypotheses H3a-H3d

Paths from	Paths to				
	Perceived TPS Success	std. error	t-value	p-value	Effect size f^2
Repair	0.450***	0.136	3.302	0.000	0.360
Internal Transparency	0.321**	0.160	2.005	0.022	0.181
Global Transparency	-0.203	0.178	1.141	0.127	0.082
Flexibility	-0.057	0.113	0.503	0.308	0.007
TPS Integration	0.263*	0.170	1.547	0.061	0.133
Mult. R^2	0.605				

Panel A of Table 6 presents path coefficients and significance levels from the PLS analysis based on bootstrapping with 5,000 drawings for the first two sets of hypotheses (H1 and H2a-d). Panel B of Table 6 presents the results for the third set of hypotheses (H3a-d) and additionally reports the effect size f^2.

***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent levels, respectively, for a one-tailed test of statistical significance.

Table 7. Results of the PLS analysis for the model optimized in SmartPLS

Panel A: Results for hypotheses H1 and H2a-H2d

Paths to	Paths from				
	TPS Integration	std. error	t-value	p-value	mult. R^2
Repair	0.351**	0.171	2.056	0.020	0.123
Internal Transparency	0.455***	0.138	3.301	0.000	0.207
Global Transparency	0.410***	0.133	3.080	0.001	0.168
Flexibility	0.311**	0.178	1.748	0.040	0.097
Perceived TPS Success	0.177	0.179	0.990	0.161	

Panel B: Results for hypotheses H3a-H3d

Paths from	Paths to				
	Perceived TPS success	std. error	t-value	p-value	Effect size f^2
Repair	0.373***	0.150	2.489	0.006	0.252
Internal Transparency	0.419***	0.162	2.590	0.005	0.219
Global Transparency	0.064	0.167	0.382	0.351	0.005
Flexibility	-0.150	0.146	1.033	0.151	0.047
TPS Integration	0.177	0.179	0.990	0.161	0.058
Mult. R^2	0.607				

Panel A of Table 7 presents path coefficients and significance levels for the model optimized in SmartPLS based on bootstrapping with 5,000 drawings for the first two sets of hypotheses (H1 and H2a-d). Panel B of Table 7 presents the results for the third set of hypotheses (H3a-d) and additionally reports the effect size f^2.

***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent levels for a one-tailed test of statistical significance, respectively.

Table 8. Results of the PLS analysis including control variables

Panel A: Results for hypotheses H1 and H2a-H2d

Paths to	Paths from				
	TPS Integration	std. error	t-value	p-value	mult. R^2
Repair	0.336**	0.172	1.957	0.025	0.113
Internal Transparency	0.388***	0.157	2.479	0.007	0.151
Global Transparency	0.246*	0.175	1.407	0.080	0.061
Flexibility	0.183	0.171	1.073	0.142	0.033
Perceived TPS Success	0.226*	0.170	1.330	0.092	

Panel B: Results for hypotheses H3a-H3d

Paths from	Paths to				
	Perceived TPS Success	std. error	t-value	p-value	Effect size f^2
Repair	0.443***	0.137	3.235	0.001	0.390
Internal Transparency	0.257*	0.195	1.318	0.094	0.113
Global Transparency	-0.177	0.165	1.072	0.142	0.065
Flexibility	-0.019	0.123	0.154	0.439	0.001
TPS Integration	0.226*	0.166	1.358	0.087	0.107
# Profit centers	0.119*	0.085	1.402	0.080	0.037
Foreign sales	0.208*	0.141	1.475	0.070	0.088
Mult. R^2	0.648				

Panel A of Table 8 presents path coefficients and significance levels from the PLS analysis based on bootstrapping with 5,000 drawings including control variables for the first two sets of hypotheses (H1 and H2a-d). Panel B of Table 8 presents the results for the third set of hypotheses (H3a-d) and additionally reports the effect size f^2.

***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent levels for a one-tailed test of statistical significance, respectively.

Table 9. Results of the PLS analysis including control variables

Panel A: Results for hypotheses H1 and H2a-H2d

Paths to	Paths from				
	TPS Integration	std. error	t-value	p-value	mult. R^2
Repair	0.336**	0.177	1.904	0.028	0.113
Internal Transparency	0.388***	0.155	2.507	0.006	0.151
Global Transparency	0.246*	0.172	1.429	0.077	0.061
Flexibility	0.183	0.169	1.080	0.140	0.033
TPS Conflicts	-0.129	0.221	0.584	0.280	

Panel B: Results for hypotheses H3a-H3d

Paths from	Paths to				
	TPS Conflicts	std. error	t-value	p-value	Effect size f^2
Repair	0.160	0.231	0.694	0.244	0.023
Internal Transparency	-0.434***	0.207	2.092	0.018	0.166
Global Transparency	0.016	0.273	0.057	0.477	0.000
Flexibility	-0.082	0.163	0.501	0.308	0.008
TPS Integration	-0.129	0.221	0.584	0.280	
Mult. R^2	0.212				

Panel A of Table 9 presents path coefficients and significance levels from the PLS analysis based on bootstrapping with 5,000 drawings (H1 and H2a-d). Panel B of Table 9 presents the results for the third set of hypotheses (H3a-d) and additionally reports the effect size f^2.

***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent levels for a one-tailed test of statistical significance, respectively.

Table 10. Results of the OLS regression

(1)	coefficient	std. error	t-value	p-value
TPS_integration	0.291**	0.139	2.090	0.044
Repair	0.411***	0.119	3.450	0.002
Internal_transparency	0.306**	0.125	2.440	0.020
Global_transparency	-0.188	0.115	-1.640	0.109
Flexibility	-0.056	0.114	-0.490	0.625
Mult. R^2	0.605			

(2)	coefficient	std. error	t-value	p-value
TPS_integration	0.530***	0.160	3.310	0.002
Mult. R^2	0.229			

(3)	coefficient	std. error	t-value	p-value
TPS_integration	0.407**	0.188	2.170	0.036
Mult. R^2	0.113			

(4)	coefficient	std. error	t-value	p-value
TPS_integration	0.451**	0.176	2.560	0.015
Mult. R^2	0.151			

(5)	coefficient	std. error	t-value	p-value
TPS_integration	0.294	0.190	1.550	0.130
Mult. R^2	0.061			

(6)	coefficient	std. error	t-value	p-value
TPS_integration	0.205	0.181	1.130	0.265
Mult. R^2	0.034			

Table 10 reports ordinary least squares coefficient estimates, standard errors, t-values, p-values and significance levels for equations (1) – (6). ***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent levels, respectively, for a two-tailed test of statistical significance. Coefficient estimates in Table 10 are specified by the following equations:

(1) $TPS\_success = \beta_1 * \text{TPS}\_integration + \beta_2 * \text{R}epair + \beta_3 * Internal\_transparency + \beta_4 * \text{G}lobal\_transparency + \beta_5 * \text{F}lexibility + \varepsilon$

(2) $TPS\_success = \beta_1 * \text{TPS}\_integration + \varepsilon$

(3) $Repair = \beta_1 * \text{TPS}\_integration + \varepsilon$

(4) $Internal\_transparency = \beta_1 * \text{TPS}\_integration + \varepsilon$

(5) $Global\_transparency = \beta_1 * \text{TPS}\_integration + \varepsilon$

(6) $Flexibility = \beta_1 * \text{TPS}\_integration + \varepsilon$

Table 11. Overview of the sample companies

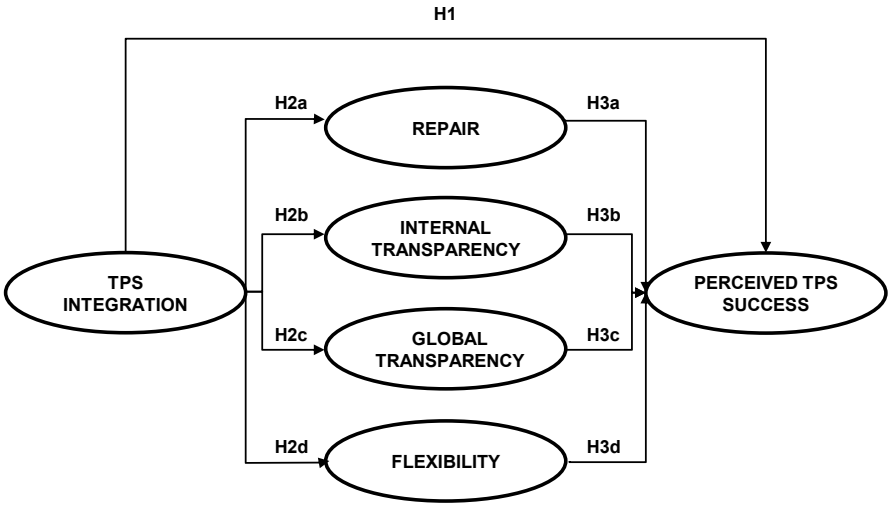
Panel A: Overview of the sample companies by organizational structure

	A	B	C
Firm size (sales in 2015)	Approx. 35 billion USD	Approx. 3 billion EUR	Approx. 70 billion EUR
Internationalization	• Revenues in various jurisdictions and currencies • Approx. 70 percent of sales generated in foreign countries	• Revenues in various jurisdictions and currencies • Approx. 80 percent of sales generated in foreign countries	• Revenues in various jurisdictions and currencies • Approx. 80 percent of sales generated in foreign countries
Organizational structure of the firm	• 4 global divisions • Each division is comprised of various business units (BU) • Each BU is comprised of various functional centers • Product selling and system selling	• 2 global divisions • No division managers, divisions are managed by the board • Each division is comprised of various functional centers	• 4 global divisions • Each business sector is comprised of various business units (BU) • Each business unit is comprised of various functional centers
Organizational structure of the value chain	• Production centers, sales and distribution centers, service centers	• Production center, service center, product center, sales center	• Entrepreneur, contract manufacturer, service provider, distributor

Panel B: Overview of the sample companies by transfer pricing system

	A	B	C
Transfer pricing guidelines	• Transfer pricing guidelines that require compliance with the arm's length principle	• Detailed and comprehensive transfer pricing guidelines	• Detailed and comprehensive transfer pricing guidelines
Transfer pricing objective	• Tax compliance as most important objective • Enhance decentralized decision-making	• Tax compliance as most important objective • Facilitate decentralized decision-making	• Tax compliance as most important objective • No distortion of decentralized decision-making
Basic structure	• Transfer prices are negotiated between the profit centers • Product selling: transfer prices are primarily determined by the production center • System selling: transfer prices are primarily determined by the sales center	• Transfer prices for the production centers are determined based on standard costs plus a markup • Transfer prices for the sales centers are determined based on the resale price minus a margin • The residual across the supply chain is allocated to the product center	• Transfer prices for the contract manufacturers and service providers are determined based on standard costs plus a markup • Transfer prices for the distributors are determined based on the resale price minus a margin • The residual across the supply chain is allocated to the entrepreneur

Figure 1: Structural equation model



Acknowledgments

We thank the following for their helpful comments and discussion: Christian Nitzl, Christian Lohmann (discussant), Dirk Simons (discussant), Hui Chen and participants at the AAA Annual Conference 2015 in Chicago, the EAA Annual Conference 2015 in Glasgow, the VHB Annual Conference 2015 in Vienna, Verein für Socialpolitik (Committee of Accounting) Annual Meeting 2015 in Mannheim, Verein für Socialpolitik (Committee of Corporate Theory and Policy) Annual Meeting 2015 in Bendorf, The Accounting Symposium 2015 in Essen, and the research seminar at the Vienna University of Economics and Business.

References

Adler, P. S., & Borys, B. (1996). Two types of bureaucracy: Enabling and coercive. *Administrative Science Quarterly, 41*(1), 61-89.

Ahrens, T., & Chapman, C. S. (2004). Accounting for flexibility and efficiency: A field study of management control systems in a restaurant chain. *Contemporary Accounting Research, 21*(2), 271-301.

Baldenius, T., Melumad, N. D., & Reichelstein, S. (2004). Integrating managerial and tax objectives in transfer pricing. *The Accounting Review, 79*(3), 591-615.

Banker, R. D., & Datar, S. M. (1992). Optimal transfer pricing under postcontract information. *Contemporary Accounting Research, 8(2), 329-352, 8*(2), 329-352.

Bouwens, J., & Steens, B. (2016). Full-cost transfer pricing and cost managment. *Journal of Management Accounting Research, 28*(3), 63-81.

Boyns, T., Edwards, J. R., & Emmanuel, C. (1999). A longitudinal study of the determinants of transfer pricing change. *Management Accounting Research, 10*(2), 89-108.

Chapman, C. S., & Kihn, L.-A. (2009). Information system integration, enabling control and performance. *Accounting, Organizations and Society, 34*(2), 151-169.

Chen, C. S., & Chen, S. (2015). Determinants and consequences of transfer pricing autonomy: An empirical investigation. *Journal of Management Accounting Research, 27*(2), 225-259.

Choe, C., & Hyde, C. E. (2007). Multinational transfer pricing, tax arbitrage and the arm's length principle. *Economic Record, 83*(263), 398-404.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Colbert, G. J., & Spicer, B. H. (1995). A multi-case investigation of a theory of the transfer pricing process. *Accounting, Organizations and Society, 20*(6), 423-456.

Cools, M. (2014). Transfer pricing: Insights from the empirical accounting literature. In D. Otley, & K. Soin (Eds.), *Management control and uncertainity* (pp. 14-29). London: Palgrave Macmillan.

Cools, M., Emmanuel, C., & Jorissen, A. (2008). Management control in the transfer pricing tax compliant multinational enterprise. *Accounting, Organizations and Society, 33*(6), 603-628.

Cools, M., & Slagmulder, R. (2009). Tax-compliant transfer pricing and responsibility accounting. *Journal of Management Accounting Research, 21*(1), 151-178.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98-104.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys. The tailored design method.* (4th ed.). Hoboken, NJ: John Wiley.

Eccles, R., G. (1985). *The transfer pricing problem: A theory of practice.* Lexington, MA: Lexington Books.

Edlin, A. S., & Reichelstein, S. (1995). Specific investment under negotiated transfer pricing: an efficiency result. *The Accounting Review, 70*, 275-291.

Eisenhardt, K., m. (1989). Building theories from case study research. *Academy of Management Review, 14*(4), 532-550.

Ernst & Young. (2001). *Transfer pricing 2001 global survey. Making informed decisions in uncertain times*. Retrieved from www.ey.com:

Ernst & Young. (2003). *Transfer pricing 2003 global survey. Practices, perceptions and trends in 22 countries*. Retrieved from www.ey.com:

Ernst & Young. (2005). *2005-2006 global transfer pricing surveys. Global transfer pricing trends, practices, and analysis, November 2005*. Retrieved from www.ey.com:

Ernst & Young. (2013). *Navigaing the choppy waters of international tax. 2013 global transfer pricing survey*. Retrieved from www.ey.com:

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50.

Free, C. (2007). Supply-chain accounting practices in the UK retail sector: Enabling or coercing collaboration? *Contemporary Accounting Research, 24*(3), 897-933.

Göx, R. F., & Schiller, U. (2007). An economic perspective on transfer pricing. In C. S. Chapman, A. G. Hopwood, & M. D. Shields (Eds.), *Handbooks of Management Accounting Research* (Vol. 2, pp. 673-695). Amsterdam: Elsevier.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: A global perspective.* (7th ed.). Upper Saddle River, NJ: Pearson.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2013). *A primer on partial least squares structural equation modeling (PLS-SEM).* Thousand Oaks, CA: Sage.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice, 19*(2), 139-152.

Hall, M. (2008). The effect of comprehensive performance measurement systems on role clarity, psychological empowerment and managerial performance. *Accounting, Organizations and Society, 33*(2-3), 141-163.

Hall, M. (2011). Do comprehensive performance measurement systems help or hinder managers' mental model development? *Management Accounting Research, 22*(2), 68-83.

Halperin, R. M., & Srinidhi, B. (1991). U.S. income tax transfer-pricing rules and resource allocation: The case of decentralized multinational firms. *The Accounting Review, 66*(1), 141-157.

Harrison, D. A., McLaughlin, M. E., & Coalter, T. M. (1996). Context, cognition, and common method variance: Psychometric and verbal protocol evidence. *Organizational Behavior and Human Decision Processes, 68*(3), 246-261.

Hartmann, F., & Slapničar, S. (2012). The fairness of performance evaluation: The role of uncertainty. *Management Accounting Research, 23*(1), 17-33.

Hirshleifer, J. (1956). On the economics of transfer pricing. *Journal of Business, 29*(3), 172-184.

Horngren, C. T., Datar, S. M., & Rajan, M. V. (2015). *Cost accounting: A managerial emphasis.* (14th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: A review of four recent studies. *Strategic Management Journal, 20*(2), 195-204.

Hyde, C. E., & Choe, C. (2005). Keeping two sets of books: The relationship between tax and incentive transfer prices. *Journal of Economics & Management Strategy, 14*(1), 165-186.

Kaplan, R., & Atkinson, A. A. (2014). *Advanced management accounting.* (3rd ed.). Harlow, Essex: Person.

Lindell, M. K., & Brandt, C. J. (2000). Climate quality and climate consensus as mediators of the relationship between organizational ante- cedents and outcomes. *Journal of Applied Psychology, 85*, 331-348.

Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology, 86*(1), 114-121.

Narayanan, V. G., & Smith, M. J. (2000). Impact of competition and taxes on responsibility center organization and transfer prices. *Contemporary Accounting Research, 17*(3), 497-529.

Nitzl, C. (2014). *Partial least squares structural equation modelling (PLS-SEM) in management accounting research: Critical analysis, advances, and future directions*. Working Paper, University of the German Federal Armed Forces – Universität der Bundeswehr München, July 22, 2014.

Nitzl, C., & Chin, W. (2016). The case of partial least squares (PLS) path modeling in managerial accounting research: Back to the future. *Journal of Management Control, forthcoming*.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory.* (3rd ed.). New York, NY: McGraw-Hill.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879-903.

Podsakoff, P. M., & Organ, D. (1986). Self-reports in organizational reserach: Problems and prospects. *Journal of Management, 12*, 69-82.

Ronen, J., & Balachandran, K. (1988). An approach to transfer pricing under uncertainty. *Journal of Accounting Research, 26*(2), 300-314.

Rossing, C. P., & Rohde, C. (2010). Overhead cost allocation changes in a transfer pricing tax compliant multinational enterprise. *Management Accounting Research, 21*(3), 199-216.

Sikka, P., & Willmott, H. (2010). The dark side of transfer pricing: its role in tax avoidance and wealth retentiveness. *Critical Perspectives on Accounting, 21*, 342-356.

Smith, M. J. (2002a). Ex ante and ex post discretion over arm's length transfer prices. *The Accounting Review, 77*(1), 161-184.

Smith, M. J. (2002b). Tax and incentive trade-offs in multinational transfer pricing. *Journal of Accounting, Auditing & Finance, 17*(2), 209-236.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics.* (4th ed.). Needham Heights, MA: Allyn and Bacon.

Van der Meer-Kooistra, J. (1994). The coordination of internal transactions: The functioning of transfer pricing systems in the organizational context. *Management Accounting Research, 5*(2), 123-152.

Van Helden, G. J., Van der Meer-Kooistra, J., & Scapens, R. W. (2001). Co-ordination of internal transactions at Hoogovens steel: Struggling with the tension between performance-oriented business units and the concept of an integrated company. *Management Accounting Research, 12*(3), 357-386.

Wouters, M., & Roijmans, D. (2011). Using prototypes to induce experimentation and knowledge integration in the development of enabling accounting information. *Contemporary Accounting Research, 28*(2), 708-736.

Wouters, M., & Wilderom, C. (2008). Developing performance measurement systems as enabling formalization: A longitudinal field study of a logistics department. *Accounting, Organizations and Society, 33*(4-5), 488-515.

Zimmermann, J. L. (2013). *Accounting for decision making and control.* (8th ed.). New York, NY: McGraw-Hill/Irwin.

2.3. Sustainability Performance and Tax Avoidance – Disentangling the Effects of Operational and Management Sustainability Performance

ABSTRACT

This paper provides an explanation for the previously mixed findings on the relationship between corporate sustainability performance (CSP) and corporate tax avoidance by disentangling sustainability performance into *operational* and *management sustainability performance*. The results from multivariate regression analyses for a sample of 4,449 U.S. firm-year observations reveal a positive relationship between management sustainability performance and tax avoidance and a negative relationship between operational sustainability performance and tax avoidance. The results are robust across the environmental and social dimensions of sustainability and various measures of tax avoidance. Overall, the study provides novel insights into the association between CSP and corporate tax avoidance.

Authors: Dr. Katrin Hummel, Benedikt Bisig

Keywords: operational sustainability performance, management sustainability performance, tax avoidance, effective tax rate, discretionary permanent book-tax differences

A version of this article was submitted to the Journal of Accounting and Public Policy.

2.3.1. Introduction

The relationship between corporate sustainability performance (CSP) and corporate tax avoidance is currently both theoretically and empirically ambiguous. From a theoretical perspective, traditional economic theories suggest that firms engage in tax avoidance and attempt to increase their sustainability performance to maximize shareholder value. Minimizing the corporate tax liability increases shareholder value, which is limited by reputational risks and the probability of detection and punishment. In this setting, firms focus on sustainability performance to build up a positive reputation, which serves as a form of risk management against corporate crises and mitigates the negative consequences of corporate crises. Therefore, traditional economic theories suggest a *positive* relationship between sustainability performance and corporate tax avoidance. By contrast, political theories including stakeholder theory and the corporate culture perspective suggest a *negative* relationship between sustainability performance and corporate tax avoidance. According to stakeholder theory, firms are responsible for a broad group of stakeholders, such as employees, the environment and society. Firms take on this responsibility by both minimizing their environmental and social externalities and paying their fair share of taxes. Similarly, the corporate culture perspective (Hill and Jones, 2001) suggests that people and groups within an organization share the same values and norms, which influence their behavior toward external stakeholders. Based on this understanding, sustainability performance and corporate tax payments complement a strong corporate culture of sustainability. Tax avoidance is thus incompatible with a high level of sustainability performance.

Taken together, the two theoretical concepts yield opposing predictions regarding the relationship between sustainability performance and corporate tax avoidance, and the few empirical studies to examine this relationship have not yet been able to clarify it. On the one hand, Davis et al. (2016) provide large-sample evidence of a positive relationship between corporate social responsibility (CSR) and corporate tax avoidance, thereby supporting traditional economic theories. On the other hand, Lanis and Richardson (2015) and Hoi et al. (2013) provide evidence of a negative relationship, which supports the applicability of political theories and the corporate culture perspective.

Trumpp et al. (2015) provide a potential explanation for these inconsistent findings. Based on a thorough review of the literature and a factor analysis of 706 firm-year observations, they develop a framework of environmental performance that distinguishes between a management

dimension and an operational dimension. The authors invite researchers to "incorporate both dimensions separately and avoid conflations" in order to derive unbiased estimations (Trumpp et al., 2015). We apply this framework to the measurement of sustainability performance and investigate the relationships between corporate tax avoidance and the management and operational dimensions separately.

The management dimension typically refers to the strategic level of sustainability performance, namely, policies, structures and processes, while the operational dimension focuses on the quantifiable environmental and social impacts of a firm. Prior research has revealed a negative relationship between operational environmental performance and a firm's environmental reputation (Cho et al., 2012), thereby suggesting that operational performance is ineffective for reputation-building and risk-management purposes. Potential reasons for such ineffectiveness may be that external stakeholders cannot easily assess a firm's operational performance because such assessments require a deeper understanding of the topic and benchmarking with a peer group. Therefore, we argue that only firms with a strong corporate culture of sustainability strive for superior operational sustainability performance, as improving the operational sustainability performance implies additional costs but not direct benefits. Consequently, we expect firms with superior operational sustainability performance to pay their fair share of taxes; thus, we also expect to find a negative relationship between operational sustainability performance and corporate tax avoidance. In addition, the management dimension of sustainability performance is more visible for a broad group of stakeholders, particularly because the management dimension also includes philanthropic activities and community engagement (Wood, 1991). Therefore, management sustainability performance is more suitable for reputation-building and risk-management purposes. Therefore, we expect traditional economic theories to be particularly applicable to management sustainability performance. Firms strive to improve their management sustainability performance based on cost-benefit considerations; thus, we expect to find a positive relationship between management sustainability performance and corporate tax avoidance.

The results from multivariate ordinary least squares (OLS) regressions with clustered standard errors at the industry-year level for a sample of 4,449 U.S. firm-year observations strongly support our reasoning. For various measures of corporate tax avoidance, namely, effective tax rates (ETRs) and discretionary book-tax differences, the results reveal a positive (negative)

relationship between management (operational) sustainability performance and corporate tax avoidance. These results hold for sustainability in general but also for the environmental and social dimensions of sustainability separately. Moreover, the results remain robust for different sample specifications and alternative measurements of tax avoidance.

Our study makes several contributions to the literature. First, this study is the first to provide a theoretical explanation that reconciles the ambiguous findings from prior empirical studies on the relationship between sustainability performance and corporate tax avoidance. Specifically, this study contributes to the empirical tax avoidance literature by providing evidence regarding potential determinants of corporate tax avoidance (Hanlon and Heitzman, 2010). Our findings suggest the integration of both management and operational sustainability performance into the models that predict corporate tax avoidance. Second, our findings also speak to the literature on CSR and sustainability (Carroll, 2015). In particular, our findings point toward the importance of distinguishing between the strategic management dimension and the outcome-oriented operational dimension of CSP, which has been previously emphasized by several researchers (Jung et al., 2001; Trumpp et al., 2015; Wood, 1991; Xie and Hayase, 2007). Our measurement of management and operational sustainability performance is closely related to the framework of Trumpp et al. (2015) and the operational sustainability performance indicators introduced by (Hummel and Schlick, 2016). While the framework of Trumpp et al. (2015) was originally been developed for environmental sustainability only, we translate this concept to overall sustainability. We thereby add to the literature on sustainability performance by providing an easily replicable measurement approach. Moreover, a separate analysis of the environmental and social dimensions of sustainability responds to several researchers' calls to study the individual sustainability dimensions separately (Hillman and Keim, 2001; Laguir et al., 2015; Rehbein et al., 2004). Third, our results show that management sustainability performance is positively and operational sustainability performance is negatively related to tax avoidance. This finding is not only statistically but also economically significant. According to our estimations, the most tax-aggressive group of firms has ETRs that are, on average, 3.5% lower than those of other firms. Each firm in this group pays, on average, $35.48 million less in annual taxes.

The remainder of this paper is structured as follows. Based on a discussion of the theoretical background and the related literature, we develop the hypotheses in the next section. Section three presents the research design, particularly the measurement of tax avoidance, the

measurement of management and operational sustainability performance, and our empirical model. In section four, we present and discuss the descriptive findings and results from multivariate regression models and additional robustness analyses. The final section concludes the paper.

2.3.2. Theoretical background

2.3.2.1. Disentangling operational and management sustainability performance

According to the triple bottom line approach, CSP is a multi-dimensional construct that comprises economic, environmental and social responsibilities (Elkington, 1997). This paper focuses on the environmental and social dimensions of sustainability. With respect to the measurement of sustainability performance, the literature agrees that sustainability performance consists of a management dimension and an operational dimension (Hummel and Schlick, 2016; Trumpp et al., 2015; Wood, 1991). On the one hand, *management* sustainability performance typically refers to the strategic level and concerns principles, policies, programs, structures and processes of sustainability (Trumpp et al., 2015; Wood, 1991). According to Wood (1991), these policies and programs include philanthropic activities and community engagement. On the other hand, *operational* sustainability performance typically focuses on the quantifiable outcomes of sustainability performance, such as the amount of greenhouse gas (GHG) emissions and waste with respect to the environment and employee turnover and injury rates with respect to the social dimension of sustainability (Hummel and Schlick, 2016; Trumpp et al., 2015). Based on factor analyses, Trumpp et al. (2015) show for the environmental dimension of sustainability that operational and management performance are distinct constructs that represent separate dimensions. Consequently, the authors caution researchers to not mix up the operational and management dimensions. Figure 1 illustrates our conceptual approach for disentangling CSP. Specifically, we distinguish between management and operational sustainability performance and between environmental and social performance.

INSERT FIGURE 1 ABOUT HERE

2.3.2.2. Traditional economic theories

Traditional economic theories suggest that firms engage in tax avoidance and attempt to increase their sustainability performance to the extent that doing so maximizes shareholder

value. With respect to tax avoidance, minimizing the amount of tax expenditures increases the firm's net income and net cash flow, which is only limited by reputational risks and the probability of detection and punishment. With respect to sustainability performance, an important mechanism through which sustainability performance enhances shareholder value is reputation and risk management. According to the risk-management argument, firms focus on sustainability performance to develop a positive sustainability reputation, which can provide "a cushion of goodwill when crises hit" (Fombrun, 1996). In other words, a high level of sustainability performance can reduce the probability of negative sustainability events and mitigate the damages that a firm encounters when negative corporate events occur (Godfrey et al., 2009; Hoi et al., 2013). Christensen (2016) provides empirical evidence of the positive role of sustainability disclosure in mitigating downside risk. In this context, CSP plays an insurance-like role against potential punishments from various stakeholders if corporate misconduct occurs. Therefore, traditional economic theories suggest a positive relationship between sustainability performance and corporate tax avoidance. Another potential explanation for a positive relationship between sustainability performance and corporate tax avoidance relates to the reasoning that corporations might be able to solve societal problems more efficiently than governments. Based on this understanding, high tax rates hamper innovation, job growth and economic development and thereby prevent firms from contributing to social welfare. Moreover, the capacity of investing in sustainability performance is higher when avoiding taxes, i.e., social welfare can be increased much more through investments in infrastructure or research and development rather than through the payment of corporate taxes.

We argue that the reasoning of traditional economic theories particularly applies to the management dimension of sustainability performance. Specifically, we argue that management sustainability performance is more visible to external stakeholders and thus more suitable for developing a positive reputation. This higher visibility stems from external stakeholders' ability to more easily assess the management sustainability performance of a firm relative to its operational sustainability performance. While the mere existence of sustainability policies, structures and processes indicates superior management sustainability performance, the assessment of operational sustainability indicators is more complicated and requires a deeper understanding of the topic and benchmarking with a peer group. Thus, firms can more easily develop a positive reputation through high management sustainability performance, which serves as insurance if corporate crises occur. This reasoning is further

supported by a study on a reputation survey provided by Shell (Bebbington et al., 2008). The authors highlight that the Shell report links reputational elements with the management dimension of sustainability (Bebbington et al., 2008). In addition, the existence of sustainability policies, structures and processes is more likely to prevent the occurrence of corporate crises than high operational sustainability performance. Philanthropic activities and community engagement are also more visible to a broad group of stakeholders and are thus more suitable for reputation building.

Therefore, we expect to find a positive relationship between management sustainability performance and corporate tax avoidance, which is formally stated in hypothesis H1a. In addition, we hypothesize that this positive relationship applies to both environmental (H1b) and social (H1c) sustainability performance.

H1a: There is a positive relationship between management sustainability performance and corporate tax avoidance.

H1b: There is a positive relationship between environmental management sustainability performance and corporate tax avoidance.

H1c: There is a positive relationship between social management sustainability performance and corporate tax avoidance.

2.3.2.3. Stakeholder theory and the corporate culture perspective

In contrast to traditional economic theories, stakeholder theory and the corporate culture perspective generally suggest a negative relationship between CSP and corporate tax avoidance. According to stakeholder theory, firms are responsible to a broad group of stakeholders, including employees, the environment and society (Freeman, 1984). By paying their fair share of taxes, firms act responsibly toward society, but they may simultaneously neglect part of the shareholder value maxim. This reasoning is in line with the existence of a "corporate culture of sustainability." According to Hill and Jones (2001), corporate culture refers to "the values and norms that are shared by people and groups in an organization and that control the way they interact with each other and with stakeholders outside the organization." Consequently, a strong corporate culture of sustainability is reflected by sustainable behavior toward all stakeholders. Tax avoidance is thus incompatible with a high level of sustainability performance. In this case, CSP and tax payments complement a strong culture of sustainability.

We argue that the corporate culture perspective particularly applies to the operational dimension of sustainability performance. Prior research has shown that operational environmental performance is negatively related to a firm's perceived environmental reputation (Cho et al., 2012).[1] Therefore, superior operational sustainability performance appears to be not suitable for developing a positive reputation, as "[operational] performance might not always be reflected in perceptions of corporate reputation" (Cho et al., 2012). External stakeholders cannot easily assess a firm's operational sustainability performance. Specifically, the assessment of quantifiable performance indicators, such as the amount of GHG emissions, energy consumption or injury rates, requires a deeper understanding of the topic and benchmarking with a peer group. Therefore, superior operational sustainability performance does not translate into a good corporate reputation. In addition, outperforming peers in terms of operational sustainability performance can be very costly. We thus argue that only firms with a strong corporate culture of sustainability invest in reducing their environmental and social impacts to levels below the common industry standards, as these reductions imply additional costs but no direct benefits. Consequently, we expect to find a negative relationship between operational sustainability performance and corporate tax avoidance, which is formally stated in hypothesis H2a. Again, we also argue that this relationship pertains to the environmental (H2b) and social (H2c) dimensions of sustainability.

H2a: There is a negative relationship between operational sustainability performance and corporate tax avoidance.

H2b: There is a negative relationship between environmental operational sustainability performance and corporate tax avoidance.

H2c: There is a negative relationship between social operational sustainability performance and corporate tax avoidance.

2.3.2.4. Empirical evidence

Our reasoning for disentangling management and operational sustainability performance might explain the inconsistent findings of prior studies regarding the relationship between sustainability performance and tax avoidance. While Davis et al. (2016) and Laguir et al. (2015) provide empirical evidence of a positive relationship between CSP and corporate tax avoidance, thereby supporting the reasoning of traditional economic theories, Hoi et al.

[1] Cho et al. (2012) measure operational environmental performance based on quantifiable indicators, "including greenhouse gas emissions, water use, solid waste disposal, acid rain emissions, and toxic waste emissions."

(2013) and Lanis and Richardson (2015) provide evidence of a negative relationship, which is in line with the corporate culture perspective.[2]

Based on a sample of 5,588 firm-year observations, Davis et al. (2016) reveal a negative (positive) relationship between firms' levels of CSR activities and tax payments (tax lobbying expenses). The level of CSR activities is measured based on data provided by the MSCI (formerly KLD) database.[3] Remarkably, this result is particularly driven by firms with high CSR levels. Similarly, Laguir et al. (2015) reveal for a sample of 83 firm-year observations a negative relationship between firms' levels of *economic* sustainability and their ETRs (as well as a positive relationship between firms' levels of *social* sustainability and their ETRs). Hoi et al. (2013) investigate a sample of U.S. firms that comprise between 4,191 and 9,147 firm-year observations.[4] Similar to Davis et al. (2016), the authors refer to the MSCI (KLD) database to construct a measure of CSR performance. However, the authors concentrate on "excessive irresponsible CSR activities" and thus include only the CSR concerns in their overall CSR measure (Hoi et al., 2013). Likewise, the authors focus on "extremely aggressive tax avoidance practices" with respect to firms' tax aggressiveness (Hoi et al., 2013). The study provides evidence for a negative and significant relationship between CSR and tax avoidance. The results are particularly applicable for a subset of firms with particularly low CSR performance, which are particularly tax aggressive. Lanis and Richardson (2015) investigate a matched sample of 434 U.S. firm-year observations. They measure tax aggressiveness as a binary variable based on self-reported tax disputes (as provided in the MSCI database). Again, the level of CSR performance is measured based on the net value of CSR strengths minus CSR concerns, as provided by the MSCI (KLD) database. The results reveal a negative and significant relationship between CSR performance and the likelihood of tax disputes. Additional separate analyses for the CSR categories indicate that this relationship particularly pertains to community relations, diversity, human rights, and product responsibility.

Taken together, except for Laguir et al. (2015), the previous empirical studies rely on data provided by the MSCI (formerly KLD) database for the measurement of CSP. The database provides binary assessments (i.e., presence versus absence) of strengths and weaknesses in

[2] Note that this paper focuses on CSP, whereas prior studies typically focus on CSR. However, the differences between the two concepts are negligible (Carroll, 2015).
[3] More precisely, the authors construct a CSR index based on the net value of CSR strengths minus CSR concerns. CSR strengths and concerns are measured in the areas of community relations, diversity, employee relations, the environment and products.
[4] The sample size varies depending on the tax avoidance measure.

sustainability-related areas, such as the environment, employee relations, and the community. However, it does not distinguish between management and operational sustainability performance. Therefore, the inconsistent findings of prior studies might relate to the mixture of management and operational sustainability performance.

2.3.3. Research design

2.3.3.1. Measuring tax avoidance

We follow Hanlon and Heitzman (2010) and broadly define corporate tax avoidance as the "reduction of explicit taxes." Note that this broad understanding of tax avoidance does not distinguish between the measures used to minimize taxes. Moreover, this broad definition captures a continuum of tax strategies that range from rather modest and legal tax planning activities to illegal tax evasion.[5] Consistent with prior research, we measure tax avoidance based on accounting data. Specifically, we rely on ETR measures, which are calculated by dividing an estimate of tax liability by a measure of before-tax profits. Notably, all measures of tax avoidance have their own limitations (for an overview see Hanlon and Heitzman, 2010). We choose ETR measures because they account for the entire continuum of tax avoidance strategies and thus allow us to gauge the *degree* of tax avoidance.[6] Moreover, prior studies on the relationship between CSR and tax avoidance have also used ETR measures, which makes our results comparable to the findings of prior studies (Davis et al., 2016; Hoi et al., 2013; Lanis and Richardson, 2012). Lanis and Richardson (2012) argue that ETRs are the most frequently used and validated proxies for tax aggressiveness in the literature and that the Australian Taxation Office considers them key indicators of tax aggressiveness.

We use a long-run cash ETR (*CASH_ETR*) and a GAAP ETR (*GAAP_ETR*) for our baseline regression models. While the numerator in the GAAP_ETR is accounting earnings, the numerator in the CASH_ETR is cash taxes paid. Each measure is affected by different tax planning strategies[7]; therefore, using both measures captures a broader spectrum of tax planning. Following Dyreng et al. (2008) and Davis et al. (2016), we apply a five-year horizon (from t to t-4) when measuring the long-run cash ETR. This type of long-run computation reduces volatility compared with year-to-year measures and eliminates a large

[5] We do not differentiate between legal and illegal tax avoidance, as determining the legality of tax avoidance based on accounting data is difficult (Hanlon and Heitzman, 2010).
[6] By contrast, the identification of firms engaged in tax shelter activities focuses on extremely aggressive tax activities at one end of the continuum.
[7] On the one hand, the GAAP ETR is affected by changes in the tax accounting accruals but not by tax deferral strategies. On the other hand, the Cash ETR is affected by tax deferral strategies but not by changes in the tax accounting accruals (Hanlon and Heitzman, 2010).

portion of the mismatch between earnings and tax payments (Hanlon and Heitzman, 2010). $CASH\_ETR_{i,t}$ is measured as the five-year average of cash taxes paid divided by pretax book income less special items for firm i in year t. Consistent with the literature, we set *CASH_ETR*=1 if it exceeds one or if the denominator is negative and cash taxes paid are positive, and we set *CASH_ETR*=0 if cash taxes paid are negative. $GAAP\_ETR_{i,t}$ is measured as the total income taxes divided by the pretax income of firm i in year t. Again, we set *GAAP_ETR*=1 if it exceeds one or if the pretax income is negative and the total income taxes are positive, and we set *GAAP_ETR*=0 if the total income taxes are negative. Note that a lower ETR reflects a higher level of corporate tax avoidance. Therefore, measuring corporate tax avoidance based on ETRs reverses the direction of our hypotheses.

2.3.3.2. Measuring operational and management sustainability performance

To measure a firm's CSP, we use data from the ASSET4 database. The ASSET4 database provides firm-level data on environmental, social and corporate governance (ESG) ratings for more than 4,300 companies worldwide (Thomson Reuters, 2012).[8] In contrast to the MSCI (KLD) database, the ASSET4 database provides not only the aggregated scoring of a firm's sustainability performance but also indicator-level information. Such granular data are necessary for the distinction between operational sustainability performance and management sustainability performance, which is an integral part of our research design.

Figure 2 provides an overview of the measurement of CSP along the management and operational dimensions as well as along the environmental and social dimensions. Our main variables of interest are management sustainability performance (*MP*), environmental management performance (*EMP*) and social management performance (*SMP*) with respect to hypotheses H1a–H1c, and operational sustainability performance (OP), environmental operational performance (*EOP*) and social operational performance (*SOP*) with respect to hypotheses H2a–H2c.

INSERT FIGURE 2 ABOUT HERE

---

[8] Data from the ASSET4 database have been increasingly used by researchers, for instance Trumpp et al. (2015), Cheng et al. (2014), and Ziegler et al. (2011).

As discussed in section 2.1, CSP is measured based on the environmental and social dimensions of sustainability. Specifically, CSP is calculated as the mean of the environmental pillar score (*ENVP*) and the social pillar score (*SOCP*), as provided by the ASSET4 database. The pillar scores reflect a firm's aggregate environmental and social performance and range between zero and 100, with lower values indicating poorer performance. The environmental pillar "examines factors including resource usage and reduction; emissions and emissions reductions; environmental activism and initiative and product or process innovation" (Thomson Reuters, 2013). The social pillar "examines factors including employment quality, health and safety issues, training, diversity, human rights, community involvement and product responsibility" (Thomson Reuters, 2013). Our distinction between operational and management sustainability follows Trumpp et al. (2015). More precisely, *EOP* and *SOP* are each measured based on an aggregation of four distinct operational performance indicators. *OP* is then measured as the arithmetic mean of *EOP* and *SOP*. *MP* (*EMP* and *SMP*) is measured as the residual from a regression of *CSP* (*ENVP* and *SOCP*) on *OP* (*EOP* and *SOP*) to capture the portion of *CSP* (*ENVP* and *SOCP*) that is not explained by *OP* (*EOP* and *SOP*). Further details on the measurement of the variables are presented below.

With respect to the operational performance dimension, *EOP* is calculated as the mean value of four indicators—energy consumption (EOP1), water withdrawal (EOP2), GHG (GHG) emissions (EOP3) and the weight of waste (EOP4)—scaled by total sales. These indicators are consistent with Trumpp et al. (2015) and Hummel and Schlick (2016).[9] Table 1 provides a detailed description of the environmental and social operational performance indicators. All EOP indicators are winsorized at the top and bottom 1% level for each year, and they are standardized by industry[10] to ensure the comparability of the performance scores across different industries and to allow for the aggregation of performance indicators into an overall score.[11] Next, the *EOP* indicators are rescaled by multiplying them by (-1) so that higher values correspond with better environmental performance. *EOP* is then calculated as the arithmetic mean of the standardized items. Thus, positive (negative) values of *EOP* can be interpreted as better (worse) than average industry performance.

[9] In contrast to Trumpp et al. (2015), we do not include hazardous waste due to the poor reporting quality of this performance indicator.
[10] For instance, GHG emissions (normalized by total sales) equal, on average, 512 for the transportation industry and 28 for the media industry. This huge discrepancy highlights the necessity of rescaling the performance data on an industry level to adequately capture the firm's environmental performance.
[11] Missing values are replaced by the mean values after standardization, i.e., by zero (mean replacement).

INSERT TABLE 1 ABOUT HERE

Similar to our procedure for measuring *EOP*, *SOP* is calculated based on the following items: employee turnover (SOP1), employee training (SOP2), the injury rate (SOP3), and the percentage of women employees (SOP4). These items are consistent with the approach of Hummel and Schlick (2016) for measuring the social dimension of sustainability performance. For interpretative reasons, we rescale employee turnover and the injury rate by multiplying them by (-1). Again, to allow for comparison across different industries, we standardize SOP3 (the injury rate) and SOP4 (the percentage of women employees) by industry and SOP1 and SOP2 on the overall level.[12] *SOP* is then calculated as the arithmetic mean of the four standardized *SOP* indicators, and *OP* is calculated as the arithmetic mean of *EOP* and *SOP*.

Trumpp et al. (2015) provide a detailed framework for measuring not only EOP but also EMP. However, our approach for measuring the management performance dimension differs from Trumpp et al. (2015) procedure for two reasons. First, the literature provides no guide for the selection of distinct management performance items for the social dimension of sustainability, and the EMP items provided by Trumpp et al. (2015) cannot be directly transferred into corresponding SMP items. Therefore, the selection of SMP indicators will suffer from considerable subjectivity. Such a potentially biased selection of SMP indicators will hamper the comparability of our results across the environmental, social and overall sustainability dimensions. Therefore, we utilize a different procedure for the measurement of management performance, which enables us to consistently derive the measures for *MP*, *EMP* and *SMP*. To measure *MP*, we regress *CSP* on *OP* according to equation (1). We then use the residuals from this regression to proxy for *MP*.

$$CSP_{i,t} = \beta_0 + \beta_1 OP_{i,t} + \epsilon_{i,t}, \quad (1)$$

[12] The summary statistics of the average indicator values per industry (untabulated) support the feasibility of this approach. Again, missing values are set to 0 after standardization.

Similarly, to measure *EMP* (*SMP*), we regress the environmental (social) pillar scores on the *EOP* (*SOP*) measures, as described in equations (2) and (3), respectively, and we use the residuals from this regression as a proxy for *EMP* (*SMP*).

$$ENVP_{i,t} = \beta_0 + \beta_1 EOP_{i,t} + \epsilon_{i,t}, \quad (2)$$

$$SOCP_{i,t} = \beta_0 + \beta_1 SOP_{i,t} + \epsilon_{i,t}, \quad (3)$$

In the robustness section of this paper (section 4.3), we further investigate the overlap between our measurement of *EMP* and the measurement of *EMP* according to Trumpp et al. (2015), and we use the latter as alternative measure for *EMP* in our regressions.

2.3.3.3. Empirical model and sample selection

The relationship between operational (management) sustainability performance and tax avoidance is analyzed based on the following regression model:

$$TA\ measure_{i,t} = \beta_0 + \beta_1 MP_{i,t} + \beta_2 OP_{i,t} + \sum \beta_k Controls_{i,t} + Year_t + Indusrty_k + \epsilon_{i,t} \quad (4)$$

The variables are summarized in Table I in the appendix. *TA measure*$_{i,t}$ proxies for our tax avoidance measures *CASH_ETR* and *GAAP_ETR*, which are described in further detail in section 3.1. *MP*$_{i,t}$ (*OP*$_{i,t}$) captures a firm's management (operational) sustainability performance as described in the previous section. Note that we estimate separate regressions for the overall sustainability dimension (i.e., including *MP*$_{i,t}$ and *OP*$_{i,t}$), the environmental dimension of sustainability (i.e., including *EMP*$_{i,t}$ and *EOP*$_{i,t}$ instead of *MP*$_{i,t}$ and *OP*$_{i,t}$), and the social dimension of sustainability (i.e., including *SMP*$_{i,t}$ and *SOP*$_{i,t}$ instead of *MP*$_{i,t}$ and *OP*$_{i,t}$).

Consistent with the literature on tax avoidance, we include a number of additional variables (*Controls*$_{i,t}$) to control for their potential influence on a firm's tax avoidance (Davis et al., 2016; Gupta and Newberry, 1997; Hoi et al., 2013). Specifically, we control for a firm's cash holdings (*CASH*), profitability (*ROA*), leverage (*LEV*), change in net operating losses (*Delta NOL*), portion of foreign income (*FI*), property, plant and equipment (*PPE*), intangible assets (*INTANG*), equity income in earnings *(EQINC),* research and development expenditures (*R&D*), size (*SIZE*), changes in sales (*Delta Sales*), the market-to-book ratio (*MB*) and selling general and administrative costs (*SG&A*). Detailed descriptions of our control variables are provided in Table I in the appendix. We winsorize all control variables at the top and bottom 1% level to minimize the influence of outliers. Whenever the dependent variable is the long-

run cash ETR, we use five-year average values not only for the measurement of *CASH_ETR* but also for *MP*, *OP* and the control variables. Additionally, we include year- and industry-fixed effects to control for time and industry-specific variance in our data.[13] The t-statistics are calculated based on standard errors clustered by industry-year.

For separate regressions on the environmental and social dimensions of sustainability, we include an additional control variable—*CSP_wo_ENVP* for the environmental dimension and *CSP_wo_SOCP* for the social dimension—to prevent omitted variable bias from excluding the other dimension of sustainability.[14] The variables are calculated as the residuals from the following regressions and thus capture the fractions of CSP that are not attributable to *ENVP* and *SOCP*, respectively:

$$CSP_{i,t} = \beta_0 + \beta_1 ENVP_{i,t} + \epsilon_{i,t}, \quad (5)$$

$$CSP_{i,t} = \beta_0 + \beta_1 SOCP_{i,t} + \epsilon_{i,t}, \quad (6)$$

To test our empirical model, we use data from two different databases. Data on sustainability performance are obtained from the Thomson Reuters ASSET4 ESG database, while data on accounting information are drawn from Standard & Poor's Compustat database. Table 2 displays the sample selection process for Panel A (*CASH_ETR* as a tax avoidance measure) and Panel B (*GAAP_ETR* as a tax avoidance measure). We start with the universe of U.S. firms with available environmental and social pillar scores (i.e., *ENVP* and *SOCP*) in the ASSET4 database for the years 2008–2015. For Panel A, variables are calculated as the five-year averages (from t to t-4). This results in 8,078 observations for Panel A and 7,283 observations for Panel B.[15] In line with prior tax avoidance research (Hoi et al., 2013; Lanis and Richardson, 2012), we exclude firms from the financial industry (GICS industry codes 4010, 4020 and 4030) and the utility industry (GICS industry code 5510). We also exclude observations with missing values for the dependent and independent variables. Our final samples consist of 4,449 firm-year observations for Panel A (*CASH_ETR*) and 4,427 firm-year observations for Panel B (*GAAP_ETR*).

---

[13] Industry classifications are based on three-digit Global Industry Classification Standard (GICS) codes.
[14] We do not include SOCP and ENVP as control variables due to multicollinearity.
[15] Note that we require items to be available for only three of the five years to calculate the long-run average ENVP and SOCP variables for Panel A, which results in a larger sample for Panel A relative to Panel B.

INSERT TABLE 2 ABOUT HERE

2.3.4. Empirical results

2.3.4.1. Descriptive statistics

Table 3 reports the summary statistics for our regression variables. On average, the ETR is 29% for the *CASH_ETR* and 31% for the *GAAP_ETR*, which falls within the range found in prior studies using U.S. data (Chen et al., 2010; Davis et al., 2016; Dyreng et al., 2008; Hoi et al., 2013; Rego, 2003). The operational sustainability performance measures (*OP*, *EOP* and *SOP*) have mean values of zero due to the standardization of the indicators. The management sustainability performance measures (*MP*, *EMP* and *SMP*) have mean values of zero due to their measurement as residuals. The median values for *OP*, *EOP* and *SOP* are equal to the mean values; those for *MP*, *EMP* and *SMP* are slightly negative; and those for *CSP_wo_EP* and *CSP_wo_SP* are slightly positive. As expected, the management sustainability performance measures (*MP*, *EMP* and *SMP*), as well as *CSP_wo_EP* and *CSP_wo_SP*, range between approximately -50 and +50, which reflects the calculation of these variables as residuals from their respective regressions. The summary statistics for the control variables are comparable to those found in prior studies (Davis et al., 2016; Hoi et al., 2013), thereby indicating no specific sample selection bias. Moreover, the summary statistics are similar across the two samples (Panel A and Panel B).

INSERT TABLE 3 ABOUT HERE

Table 4 reports the Pearson correlation coefficients and their respective p-values (in parentheses) for Panel A (except for the correlations between *GAAP_ETR* and other variables). As expected, our two measures of tax avoidance, *CASH_ETR* and *GAAP_ETR*, are significantly positively correlated. With regard to our sustainability performance variables (i.e., variables 3–10), we obtain three important findings. First, every operational sustainability performance measure (*OP*, *EOP* and *SOP*) has a correlation of zero with its respective management sustainability performance measure (*MP*, *EMP* and *SMP*), which reflects the construction of the management performance variables as residuals from the regressions (1), (2), and (3), respectively. Second, the operational performance measures (*OP*,

EOP, and *SOP*) and the management performance measures (*MP*, *EMP*, and *SMP*) are significantly positively associated with one another at a statistical significance level of 0.000. Further, we find a positive relation between *EOP* and *SOP* (at a significance level of 0.000), indicating that *EOP* and *SOP* are positively associated. In other words, firms with good SOP tend to also outperform their benchmark group in terms of *EOP*. Thus, our approach of aggregating *SOP* and *EOP* into an overall operational sustainability performance measure appears reasonable. With respect to our hypotheses, we find that both ETR measures are significantly negatively associated with management sustainability performance measures (*MP*, *EMP* and *SMP*), while the correlations between the ETR measures and the operational sustainability performance measures (*OP*, *EOP* and *SOP*) are positive and mostly significant. Although we do not control for potential confounding effects from the other independent variables, these findings provide some preliminary evidence in support of our hypotheses.

INSERT TABLE 4 ABOUT HERE

2.3.4.2. Findings from regression analyses

Table 5 reports the results from OLS regressions with clustered standard errors at the industry-year level for equation (4). Columns (1) to (3) display the results for *CASH_ETR* as the dependent variable (Panel A), and columns (4) to (6) display the results for *GAAP_ETR* as the dependent variable (Panel B). Moreover, columns (1) and (4) present the results for the overall dimension of sustainability; columns (2) and (5) present those for the environmental dimension of sustainability; and columns (3) and (6) present those for the social dimension of sustainability. Note that lower ETRs correspond with higher tax avoidance.

With respect to the first set of hypotheses (H1a to H1c), the results reveal negative and significant coefficients for *MP*, *EMP*, and *SMP* for both ETR measures. These findings support hypotheses H1a, H1b, and H1c, indicating a substitutive relationship between management sustainability performance (*EMP* and *SMP*) and corporate tax payments. Therefore, firms with higher management sustainability performance (*EMP* and *SMP*) have significantly lower ETRs and are thus more tax avoidant. This finding supports the reasoning of traditional economic theories that firms aim to increase their sustainability performance

and engage in tax avoidance based on cost-benefit considerations. Specifically, firms with high sustainability management performance build up a positive sustainability reputation that acts as insurance if negative corporate events occur. This positive sustainability reputation might protect them from the downside risk in cases of critical public tax debates; therefore, firms might avoid corporate taxes more aggressively. This reasoning is further enhanced, as the measure also entails corporate philanthropic activities, which are particularly useful for developing a positive reputation, to some extent. The magnitude of the coefficient for *MP* (-0.0008) indicates that a 1-unit increase (e.g., from 4 to 5) in management sustainability performance is associated with a 0.8 percentage point decrease in the long-run cash ETR (e.g., from 0.320 to 0.312).

INSERT TABLE 5 ABOUT HERE

With respect to the second set of hypotheses (H2a to H2c), the results reveal positive and significant coefficients for *OP*, *EOP*, and *SOP* for both ETR measures. Firms with higher operational sustainability performance have higher ETRs and are less tax avoidant. This finding supports hypotheses H2a to H2c and thus the reasoning derived from stakeholder theory and the corporate culture perspective of the firm. Firms with a strong corporate culture of sustainability act more responsibly toward all stakeholders. These firms have lower environmental and social externalities and pay higher taxes. In contrast to management sustainability performance, the operational sustainability performance is less visible to external stakeholders and is thus more likely to reflect the genuine sustainability culture of the firm. Because the operational performance indicators are standardized, the magnitude of the coefficients cannot be reasonably interpreted.

Our findings on the control variables are mostly consistent with Davis et al. (2016) and Hoi et al. (2013). Specifically, we find that larger and more profitable firms have lower ETRs, probably due to more resources in tax planning and access to a broader scope of tax planning activities. In addition, foreign income and intangibles are negatively associated with the ETR measures, which supports the reasoning that firms can reduce their tax payments through

foreign activities and intangibles. We find a positive relationship between leverage and the ETR, which suggests that firms with more debt have higher tax rates.

Next, we are interested in the economic relevance of our results. Therefore, we replace *MP* with a dummy variable *pos_MP*, which equals 1 if a firm has a positive *MP* and 0 otherwise. We re-run the regressions according to equation (4) with *pos_MP* instead of *MP* as our main variable of interest. The results are displayed in Table 6. The statistically significant coefficient on *pos_MP* equals -0.0187, which indicates that the cash ETR of firms with positive management sustainability performance is, on average, 1.9% lower than that of firms with negative management sustainability performance. We apply the same procedure for *OP* by introducing a dummy variable for negative operational performance, which equals 1 if *OP* is negative and 0 otherwise. The significantly negative coefficient of -0.0200 on *neg_OP* (regression 2) indicates that the long-run cash ETR of firms with negative operational sustainability performance is, on average, 2% lower (compared with that of firms with non-negative operational sustainability performance). Given these results, we create the interaction term *pos_MP x neg_OP* to capture firms that we expect to participate in the most tax avoidance in given years. The statistically significant negative coefficient on *pos_MP x neg_OP* (regression 3) of -0.0347 indicates that the cash ETR of firms with both positive management sustainability performance and negative operational sustainability performance is, on average, 3.5% lower. In dollars, the ETR difference between the two groups equals, on average, $35.48 million in taxes paid annually (calculated based on a mean pretax income of $1.023 billion), which is higher than the tax savings reported by Hoi et al. (2013) for firms with highly negative CSR activities. Note that both the magnitude of the estimated coefficient and the average pretax income are higher in our study. This finding suggests that outperforming the benchmark group in terms of management sustainability performance and having lower operational sustainability performance than the benchmark average will lead to economically relevant lower tax payments.

INSERT TABLE 6 ABOUT HERE

2.3.4.3. Robustness

We perform a number of additional analyses to test the robustness of our main findings. First, we investigate whether our results are driven by the sample selection, particularly the exclusion of firms from the financial or utilities industry. This procedure results in 5,024 observations for Panel A and 4,951 observations for Panel B. All variables are calculated according to the procedures described in the research design section. The results from the multivariate OLS regressions with standard errors clustered at the industry-year are similar to our baseline specifications (untabulated). In particular, we again obtain negative and significant coefficients for *MP*, *EMP* and *SMP* and positive and significant coefficients for *OP*, *EOP* and *SOP*. However, the coefficients for *EOP* and *SMP* become insignificant in Panel A. Next, we investigate whether the handling of loss observations drives our main findings. In the baseline model, we include loss observations, which affect the measurement of our dependent variables (which take a value of 1 if the pretax book income is negative and income taxes are positive) and our control variables, particularly *ROA* and *FI*. Therefore, we exclude firm-year observations with negative pretax book income, which results in a reduced sample of 3,847 observations for Panel A (*CASH_ETR*) and 3,834 observations for Panel B (*GAAP_ETR*). The results (untabulated) primarily remain unchanged. In particular, the coefficients for *OP*, *MP*, *EMP*, *EOP* and *SOP* are similar to our main findings with respect to sign, magnitude and statistical significance. The coefficients for *SMP* remain negative, but they become insignificant at common significance levels. Taken together, these findings suggest that our results are not predominantly driven by sample selection. Next, one could argue that our results may be biased due to omitted variables, particularly corporate governance characteristics. Specifically, prior studies report empirical evidence of a link between corporate governance characteristics and tax avoidance (Armstrong et al., 2014; Desai and Dharmapala, 2006; Gaertner, 2014). Therefore, we include an additional control variable that proxies for the economic and corporate governance dimensions of sustainability (similar to Davis et al., 2016). The construction of this variable is similar to the construction of *CSP_wo_ENVP* and *CSP_wo_SOCP* (equations (5) and (6), respectively). However, we now regress the ASSET4 aggregate score that captures a firm's sustainability performance in the economic, environmental, social and corporate governance dimensions on *CSP*, *ENVP*, and *SOCP* and use the residuals from these regressions. *ESG_wo_CSP*, *ESG_wo_ENVP*, and *ESG_wo_SOCP* thus capture the portion of the ESG rating that is not driven by *CSP*, *ENVP*, and *SOCP*. We re-run our baseline regressions and include these additional control variables, i.e., *ESG_wo_CSP* for the overall sustainability dimension, *ESG_wo_ENVP* for the

environmental dimension, and *ESG_wo_SOCP* for the social dimension. The untabulated results for our main variables of interest (*MP*, *OP*, *EMP*, *EOP*, *SMP* and *SOP*) remain unchanged in terms of sign, magnitude and significance. The new control variables are negative and significant for Panel A (*CASH_ETR*) and negative and insignificant for Panel B (*GAAP_ETR*). Taken together, the findings of this additional analysis reveal that our results are not mainly driven by omitted variable bias.

Second, we account for the rich literature on tax avoidance, which provides a variety of measures for tax avoidance (Hanlon and Heitzman, 2010). In our baseline model, we choose ETRs as measures of tax avoidance because they focus on the entire continuum of tax avoidance strategies and because they are widely used measures in the literature on tax avoidance (Hanlon and Heitzman, 2010). Another approach for measuring tax avoidance follows the traditional approaches for measuring earnings management (e.g., Jones, 1991) and estimates the "discretionary" or "abnormal" portion of book-tax differences. Compared with the traditional ETR measures, such a measure can "capture items that alter, and in particular reduce, the firm's GAAP effective tax rate which then raises bottom-line earnings, all else constant" (Hanlon and Heitzman, 2010). Therefore, we follow Frank et al. (2009) and use the discretionary portion of the permanent book-tax difference (*DTAX*) as an alternative measure of tax avoidance. The permanent book-tax difference (*PERMDIFF*) is the difference between the effective and statutory tax rates multiplied by pretax income. *PERMDIFF* is regressed on typical determinants such as the amount of intangible assets or the change in the tax loss carryforward (see Table I in the appendix for details). The residuals from this regression are used to proxy for *DTAX*. In contrast to the traditional ETR measures, *DTAX* aims at removing "unintended" tax avoidance and capturing the portion of truly intentional tax avoidance. The results from our regression models with *DTAX* as an alternative measure of tax avoidance are displayed in Panel A of Table 7.[16] Note that higher *DTAX* values reflect a higher portion of discretionary book-tax differences and thus a higher level of tax avoidance. The signs of the coefficients for MP and OP are thus reversed compared with the results in Table 5. Similar to our baseline results, the findings unanimously support our hypotheses. Firms with higher management sustainability performance have higher discretionary book-tax differences, while firms with higher operational sustainability performance have lower discretionary book-tax differences. Taken together, these findings provide evidence that our results are not solely driven by the measurement of corporate tax avoidance.

[16] Similar to Hoi et al. (2013), we include the lagged DTAX as an additional control variable in our regressions.

INSERT TABLE 7 ABOUT HERE

Third, another concern with our main findings relates to the measurement of our sustainability performance variables. Specifically, only the measurements of *EOP* and *SOP* reflect those used in the literature (Hummel and Schlick, 2016; Trumpp et al., 2015), whereas our approach to measuring *MP*, *EMP* and *SMP* is unique in this research context. First, we alternatively measure MP_{mean} as the mean of EMP and SMP and re-run our baseline regressions with MP_{mean} instead of MP. The results remain robust for both the *CASH_ETR* and *GAAP_ETR* (untabulated). Next, because Trumpp et al. (2015) provide a theoretically sound and empirically assessed framework for the measurement of *EMP*, we alternatively measure *EMP* according to this framework (EMP_{TR}). Depending on the samples used (Panel A or Panel B, respectively), the correlations between *EMP* and EMP_{TR} range between 0.88 and 0.90 (p-values of 0.0000), thereby indicating very high correlations between the two measures. Therefore, the two measures are likely to measure similar aspects of sustainability performance. Finally, we re-run our baseline regressions for the environmental dimension of sustainability with EMP_{TR} instead of *EMP* and the dependent variables *CASH_ETR*, *GAAP_ETR*, and *DTAX*. The results are displayed in Panel B of Table 7. Except for the coefficient for EOP in regression (4), the results remain robust for both *EOP* and *EMP*. The results from this additional analysis help validate our measurements of *MP*, *EMP*, and *SMP* and show that our results are not particularly driven by our measurement approach.

Taken together, the results from these additional analyses indicate that our results are robust to different sample specifications, measures of corporate tax avoidance, and CSP measures.

2.3.5. Conclusions

This paper investigates the relationship between CSP and corporate tax avoidance by distinguishing between management and operational sustainability performance. The management dimension refers to a firm's principles, policies, programs, structures and processes of sustainability performance, whereas the operational dimension refers to a firm's environmental and social impacts. Consistent with traditional economic theories, we hypothesize that there is a positive relationship between management sustainability

performance and corporate tax avoidance. By contrast, in line with stakeholder theory and the corporate culture perspective, we expect to find a negative relationship between operational sustainability performance and corporate tax avoidance. The results from the OLS regressions of a sample of 4,449 U.S. firm-year observations support this reasoning. The results are robust to different measures of corporate tax avoidance, namely, a long-run cash ETR, a GAAP ETR, and a discretionary book-tax difference, and they pertain to both overall sustainability and the environmental and social dimensions separately.

The practical implications of this research relate to two areas. First, considering the significant relationships between CSP and tax avoidance, stakeholders need additional information about a firm's tax payments and corporate tax strategy to derive a balanced picture of its behavior toward society. Therefore, a first step may be to integrate this information into common sustainability disclosure guidelines, such as the GRI guidelines and the United Nations Global Compact (UNGC). Second, our findings reveal that the relationship between CSP and corporate tax avoidance depends on whether the management dimension or the operational dimension of sustainability is considered. Stakeholders need to be aware of these differences when assessing a firm's CSP.

As with all papers, the results of this paper are also subject to certain limitations. First, the generalizability of our findings depends on our sample. In particular, this study focuses solely on U.S. firms. While the focus on U.S. firms might enhance the comparability of our findings because most prior studies examine U.S. samples, the results may not hold in a European context. In fact, Matten and Moon (2008) argue that the understanding of CSR differs significantly between coordinated (i.e., European) and liberal (i.e., U.S.) economies. Second, our measures of corporate tax avoidance are obtained from accounting data, which may yield erroneous estimations of taxable income. While the use of tax returns enables researchers to more accurately estimate the firm's taxable income, these data are seldom publicly available, and they are often difficult to match with financial statement data (Hanlon and Heitzman, 2010). We thus rely on multiple measures that are commonly used in the literature to capture tax avoidance. The robustness of our findings across these different measures dispels potential concerns regarding the validity of our measures. A third caveat of this study relates to the measurement of operational and management sustainability performance. While we closely follow the existing literature in constructing our operational performance measure, the literature on the measurement of the management dimension is limited; therefore, we develop

our own measurement approach. The substantial overlap between our variable EMP and the variable EMP_{TR}, which is calculated according to Trumpp et al. (2015), provides contributes to the validity of our measurement approach.

These limitations provide opportunities for future research. First, additional studies could investigate whether our findings hold in a European context, where a more implicit understanding of CSR persists. Second, future research could empirically validate our measures of management sustainability performance and social management performance or develop and validate alternative measures based on the data items drawn from the ASSET4 database. In addition, as our results indicate that management sustainability performance is more related to a traditional economic understanding of corporate sustainability while operational sustainability performance is more related to the corporate culture of a firm, researchers might be interested in determining whether these differences also hold for other areas of research. Third, our results reveal significant relationships between CSP and corporate tax avoidance. While Davis et al. (2016) provide some anecdotal evidence of how firms discuss tax payments in their sustainability reports, further research is needed to complement our findings from a corporate sustainability disclosure perspective.

Figure 1: Conceptual approach to measuring corporate sustainability performance

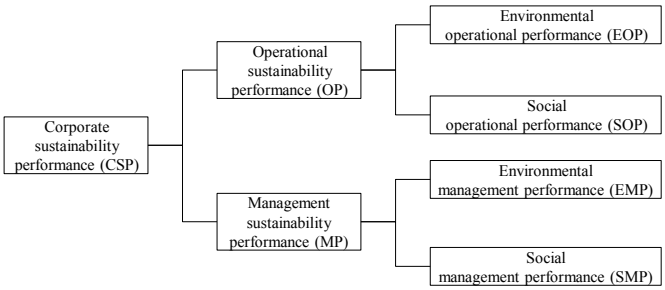


Figure 2: Measurement of operational and management sustainability performance

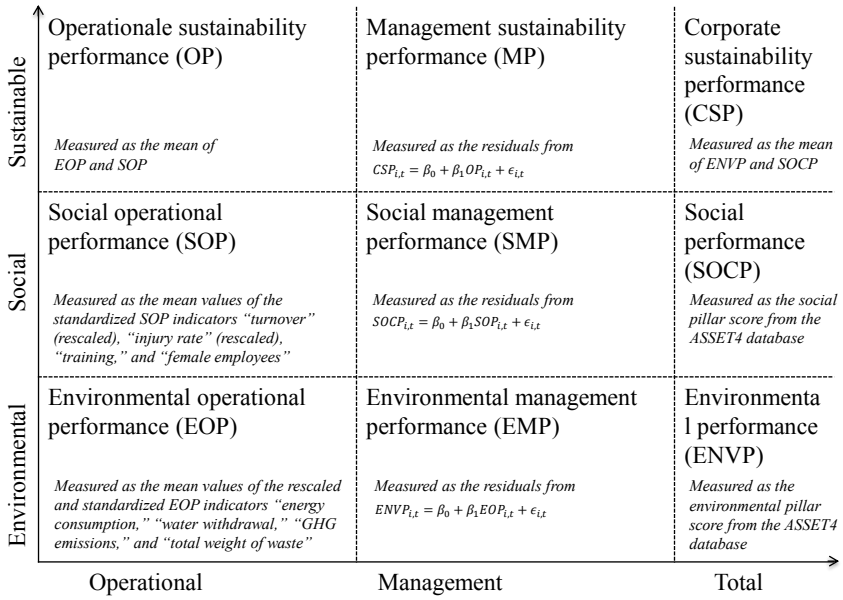


Table 1: Measurement of environmental and social operational performance

Code	ASSET4 item	Performance indicator	Measurement
Environmental dimension			
EOP1	ENRRDP033	Energy consumption	(Total direct and indirect energy consumption)/total sales
EOP2	ENRRDP054	Water withdrawal	(Total water withdrawal in cubic meters)/total sales
EOP3	ENERDP023	Greenhouse gas emissions	(Total CO_2 and CO_2 equivalent emissions in tonnes)/total sales
EOP4	ENERDP045	Total weight of waste	(Total amount of waste produced in tonnes)/total sales
Social dimension			
SOP1	SOEQDP034	Turnover	Percentage of employee turnover
SOP2	SOHSDP024	Injury rate	Total number of injuries and fatalities including no-lost-time injuries relative to one million hours worked
SOP3	SOTDDP018	Training	Average hours of training per year per employee
SOP4	SODODP017	Female employees	Percentage of female employees

Table 2. Sample selection

	Panel A: CASH_ETR	Panel B: GAAP_ETR
Firm-year observations with available ENVP and SOCP scores (ASSET4 ESG Database)	8,078	7,283
-Financial or utility companies	-1,450	-1,422
-Missing values for dependent variable	-1,607	-184
-Missing values for independent variables	-572	-1,250
Resulting firm-year observations	**4,449**	**4,427**

This table reports the sample selection for the baseline regression models.

Table 3. Summary statistics

	Panel A: CASH_ETR						Panel B: GAAP_ETR					
Variable	n	Mean	SD	Min	P50	Max	n	Mean	SD	Min	P50	Max
Dependent variables												
CASH_ETR	4,449	0.29	0.23	0.00	0.25	1.00						
GAAP_ETR							4,427	0.31	0.23	0.00	0.30	1.00
Main variables of interest												
MP	4,449	0.00	22.61	-46.66	-3.15	47.03	4,427	0.00	24.36	-49.05	-1.65	47.91
OP	4,449	0.00	0.14	-1.61	0.00	0.62	4,427	0.00	0.20	-1.75	0.00	0.78
EOP	4,449	0.00	0.26	-3.05	0.00	0.92	4,427	0.00	0.33	-3.29	0.00	1.02
EMP	4,449	0.00	29.58	-34.91	-8.14	53.43	4,427	0.00	32.03	-38.81	-6.58	48.20
SOP	4,449	0.00	0.12	-1.04	0.00	0.85	4,427	0.00	0.20	-2.00	0.00	1.20
SMP	4,449	0.00	26.23	-42.71	-2.80	52.15	4,427	0.00	27.68	-46.21	-1.44	57.47
Control variables												
CSP_wo_EP	4,449	0.00	14.07	-47.17	0.66	41.58	4,427	0.00	16.45	-57.67	1.27	55.07
CSP_wo_SP	4,449	0.00	8.62	-29.35	0.07	27.21	4,427	0.00	11.04	-33.68	0.85	32.94
CASH	4,449	0.17	0.17	0.00	0.11	0.87	4,427	0.16	0.17	0.00	0.11	0.84
ROA	4,449	0.10	0.10	-0.17	0.09	0.44	4,427	0.09	0.11	-0.29	0.09	0.46
LEV	4,449	0.25	0.19	0.00	0.21	0.99	4,427	0.25	0.21	0.00	0.22	1.14
Delta NOL	4,449	0.01	0.03	-0.07	0.00	0.15	4,427	0.01	0.04	-0.13	0.00	0.24
FI	4,449	0.38	1.17	-2.75	0.13	8.75	4,427	0.30	0.86	-3.07	0.11	5.11
PPE	4,449	0.30	0.25	0.01	0.22	1.10	4,427	0.29	0.25	0.00	0.21	1.08
INTANG	4,449	0.26	0.23	0.00	0.20	0.95	4,427	0.25	0.24	0.00	0.19	1.17
EQINC	4,449	0.00	0.00	-0.01	0.00	0.03	4,427	0.00	0.01	-0.01	0.00	0.03
R&D	4,449	0.03	0.05	0.00	0.00	0.26	4,427	0.03	0.05	0.00	0.00	0.25
SIZE	4,449	8.53	1.16	5.99	8.43	11.61	4,427	8.66	1.17	6.17	8.54	11.96
Delta SALE	4,449	0.08	0.12	-0.17	0.05	0.59	4,427	0.05	0.17	-0.55	0.04	0.67
MB	4,449	3.55	5.51	-19.31	2.69	35.98	4,427	3.40	4.92	-16.04	2.50	31.02
SG&A	4,449	0.22	0.19	0.00	0.17	0.85	4,427	0.21	0.18	0.00	0.16	0.81

This table reports summary statistics for the regression variables for the sample used in the baseline regression model. The values reported in Panel A are the mean values over a five-year period. All control variables are winsorized at the top and bottom 1% level (except for binary variables).

Table 4. Pearson correlation coefficients

		1)	2)	3)	4)	5)	6)	7)	8)	9)	10)	11)	12)	13)	14)	15)	16)	17)	18)	19)	20)	21)	22)	23)
1)	CASH_ETR	1.000																						
2)	GAAP_ETR	0.146	1.000																					
		(0.000)																						
3)	MP	-0.185	-0.071	1.000																				
		(0.000)	(0.000)																					
4)	OP	0.020	0.060	0.000	1.000																			
		(0.190)	(0.000)	(1.000)																				
5)	EOP	0.006	0.040	-0.023	0.862	1.000																		
		(0.691)	(0.008)	(0.122)	(0.000)																			
6)	EMP	-0.101	-0.062	0.782	0.016	0.000	1.000																	
		(0.000)	(0.000)	(0.000)	(0.299)	(1.000)																		
7)	SOP	0.025	0.041	0.041	0.521	0.086	0.029	1.000																
		(0.095)	(0.006)	(0.006)	(0.000)	(0.000)	(0.056)																	
8)	SMP	-0.114	-0.062	0.924	-0.038	-0.046	0.831	0.000	1.000															
		(0.000)	(0.000)	(0.000)	(0.012)	(0.002)	(0.000)	(1.000)																
9)	CSP_wo_EP	-0.170	-0.034	0.622	0.030	0.019	0.001	0.036	0.438	1.000														
		(0.000)	(0.025)	(0.000)	(0.046)	(0.205)	(0.957)	(0.016)	(0.000)															
10)	CSP_wo_SP	-0.212	-0.041	0.381	0.042	0.048	0.037	0.004	-0.000	0.569	1.000													
		(0.000)	(0.007)	(0.000)	(0.005)	(0.001)	(0.015)	(0.808)	(0.991)	(0.000)														
11)	CASH	0.030	-0.050	-0.142	0.077	0.079	-0.119	0.027	-0.136	-0.074	-0.045	1.000												
		(0.044)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.075)	(0.000)	(0.000)	(0.002)													
12)	ROA	-0.294	-0.028	0.150	0.073	0.071	-0.034	0.028	0.040	0.287	0.293	0.165	1.000											
		(0.000)	(0.067)	(0.000)	(0.000)	(0.000)	(0.024)	(0.059)	(0.007)	(0.000)	(0.000)	(0.000)												
13)	LEV	0.122	0.037	-0.177	-0.028	-0.023	-0.089	-0.018	-0.106	-0.174	-0.205	-0.288	-0.193	1.000										
		(0.000)	(0.013)	(0.000)	(0.059)	(0.125)	(0.000)	(0.239)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)											
14)	Delta_NOL	0.218	0.078	-0.079	0.014	0.008	-0.018	0.020	-0.042	-0.104	-0.107	0.146	-0.204	0.063	1.000									
		(0.000)	(0.000)	(0.000)	(0.347)	(0.579)	(0.241)	(0.185)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)										
15)	FI	-0.092	-0.058	0.155	-0.004	-0.013	0.135	0.020	0.132	0.079	0.084	0.052	0.027	-0.065	-0.012	1.000								
		(0.000)	(0.000)	(0.000)	(0.789)	(0.385)	(0.000)	(0.183)	(0.000)	(0.000)	(0.000)	(0.001)	(0.068)	(0.000)	(0.406)									
16)	PPE	0.079	-0.006	-0.035	-0.142	-0.147	-0.043	-0.026	-0.047	-0.011	0.023	-0.286	-0.050	0.201	0.034	-0.060	1.000							
		(0.000)	(0.707)	(0.019)	(0.000)	(0.000)	(0.004)	(0.086)	(0.002)	(0.480)	(0.128)	(0.000)	(0.001)	(0.000)	(0.023)	(0.000)								
17)	INTANG	-0.061	0.052	-0.006	0.059	0.065	-0.049	-0.003	-0.030	0.055	0.056	-0.156	-0.072	0.186	0.017	0.012	-0.428	1.000						
		(0.000)	(0.001)	(0.671)	(0.000)	(0.000)	(0.001)	(0.851)	(0.046)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.264)	(0.420)	(0.000)							
18)	EQINC	-0.069	-0.019	0.144	-0.022	-0.063	0.142	0.080	0.133	0.050	0.047	-0.127	0.031	0.013	-0.055	0.007	0.073	-0.031	1.000					
		(0.000)	(0.209)	(0.000)	(0.142)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)	(0.039)	(0.387)	(0.000)	(0.650)	(0.000)	(0.041)						
19)	R&D	0.028	-0.047	0.023	0.071	0.059	0.062	0.047	0.039	-0.037	-0.039	0.604	-0.072	-0.207	0.152	0.079	-0.289	0.024	-0.086	1.000				
		(0.060)	(0.002)	(0.129)	(0.000)	(0.000)	(0.000)	(0.002)	(0.010)	(0.012)	(0.010)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.114)	(0.000)					
20)	SIZE	-0.070	-0.053	0.493	0.001	-0.014	0.511	0.029	0.524	0.148	0.019	-0.325	-0.161	0.104	-0.050	0.047	0.099	0.087	0.167	-0.193	1.000			
		(0.000)	(0.000)	(0.000)	(0.965)	(0.346)	(0.000)	(0.056)	(0.000)	(0.000)	(0.195)	(0.000)	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)				
21)	Delta_SALE	-0.052	0.078	-0.087	0.044	0.055	-0.180	-0.002	-0.153	0.089	0.142	0.267	0.358	-0.095	0.061	0.033	0.033	0.053	-0.058	0.069	-0.197	1.000		
		(0.001)	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)	(0.912)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.030)	(0.027)	(0.000)	(0.000)	(0.000)	(0.000)			
22)	MB	-0.036	-0.002	0.042	0.056	0.046	0.031	0.038	0.032	0.031	0.028	0.191	0.184	-0.060	0.023	0.007	-0.052	-0.043	-0.012	0.137	-0.069	0.160	1.000	
		(0.015)	(0.894)	(0.005)	(0.000)	(0.002)	(0.037)	(0.011)	(0.033)	(0.042)	(0.059)	(0.000)	(0.000)	(0.000)	(0.129)	(0.655)	(0.001)	(0.004)	(0.433)	(0.000)	(0.000)	(0.000)		
23)	SG&A	0.032	0.086	0.025	0.057	0.078	-0.020	-0.020	-0.004	0.070	0.078	0.302	0.267	-0.234	0.046	0.002	-0.267	-0.018	-0.163	0.260	-0.333	0.294	0.180	1.000
		(0.035)	(0.000)	(0.096)	(0.000)	(0.000)	(0.176)	(0.192)	(0.775)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.888)	(0.000)	(0.219)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

This table reports the Pearson correlation coefficients for the regression variables. p-values (for a two-tailed test of statistical significance) are reported in brackets. The correlations among independent variables are calculated for Panel A; the correlations between independent variables and the CASH_ETR are calculated based on Panel A; and the correlations between independent variables and GAAP_ETR are calculated based on Panel B.

Table 5. Multivariate regressions

	Panel A			*Panel B*		
	(1)	(2)	(3)	(4)	(5)	(6)
Variables	CASH_ETR	CASH_ETR	CASH_ETR	GAAP_ETR	GAAP_ETR	GAAP_ETR
MP	-0.0008***			-0.0005***		
	(-4.0081)			(-3.1667)		
OP	0.0692***			0.0657***		
	(3.7183)			(3.6854)		
EMP		-0.0005**			-0.0003*	
		(-2.5894)			(-1.9541)	
EOP		0.0204*			0.0289***	
		(1.7400)			(2.7179)	
SMP			-0.0003*			-0.0003**
			(-1.7050)			(-2.0469)
SOP			0.0737***			0.0484***
			(3.2545)			(2.9521)
CSP_wo_ENVP		-0.0009***			-0.0005**	
		(-3.7772)			(-2.4594)	
CSP_wo_SOCP			-0.0028***			-0.0012***
			(-6.5848)			(-3.2343)
CASH	0.1174***	0.1171***	0.1166***	-0.1142***	-0.1144***	-0.1099***
	(3.0815)	(3.0850)	(3.1666)	(-3.8214)	(-3.8235)	(-3.6879)
ROA	-0.6675***	-0.6618***	-0.6295***	-0.0964	-0.0933	-0.0760
	(-10.4570)	(-10.5335)	(-10.4383)	(-1.4302)	(-1.3996)	(-1.1273)
LEV	0.0856***	0.0859***	0.0747***	0.0199	0.0202	0.0171
	(3.6497)	(3.6699)	(3.2473)	(0.9304)	(0.9496)	(0.7984)
Delta NOL	1.2492***	1.2518***	1.2172***	0.1962	0.1974	0.1904
	(6.7025)	(6.6854)	(6.4951)	(1.3457)	(1.3562)	(1.3037)
FI	-0.0126***	-0.0127***	-0.0123***	-0.0175***	-0.0173***	-0.0176***
	(-4.1353)	(-4.1710)	(-4.0018)	(-2.8875)	(-2.8425)	(-2.9186)
PPE	0.0186	0.0128	0.0219	-0.0200	-0.0243	-0.0259
	(0.5402)	(0.3707)	(0.6534)	(-0.8376)	(-1.0292)	(-1.0938)
INTANG	-0.0681***	-0.0678***	-0.0513**	0.0004	-0.0005	0.0077
	(-2.7461)	(-2.7193)	(-2.0260)	(0.0176)	(-0.0263)	(0.3646)
EQINC	-1.6982**	-1.6910**	-1.8829**	-0.2022	-0.1387	-0.3220
	(-2.2836)	(-2.3198)	(-2.5459)	(-0.3003)	(-0.2031)	(-0.4720)
R&D	-0.1761	-0.1698	-0.1842	0.0494	0.0564	0.0544
	(-1.4663)	(-1.4047)	(-1.5286)	(0.3286)	(0.3728)	(0.3613)
SIZE	-0.0075*	-0.0078*	-0.0111***	-0.0073*	-0.0071*	-0.0082**
	(-1.9537)	(-1.8840)	(-2.7090)	(-1.8003)	(-1.6597)	(-2.0110)
Delta SALE	-0.0349	-0.0329	-0.0248	0.0478*	0.0478*	0.0522*
	(-1.0284)	(-0.9762)	(-0.7503)	(1.7575)	(1.7424)	(1.9109)
MB	-0.0004	-0.0004	-0.0005	-0.0006	-0.0006	-0.0006
	(-0.7834)	(-0.7129)	(-0.8737)	(-0.9099)	(-0.8529)	(-0.8529)
SG&A	0.1192***	0.1190***	0.1156***	0.0909***	0.0912***	0.0908***
	(4.0784)	(4.0629)	(4.1654)	(3.3180)	(3.2783)	(3.2578)
Constant	0.4283***	0.4313***	0.4634***	0.4555***	0.4538***	0.4581***
	(9.2307)	(9.0522)	(9.8926)	(8.5762)	(8.3390)	(8.6158)
Observations	4,449	4,449	4,449	4,427	4,427	4,427
Adjusted R-squared	0.1719	0.1706	0.1773	0.0521	0.0503	0.0517
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes	Yes	Yes	Yes
F	26.287	24.803	20.647	7.4318	7.0768	7.2186

*, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively (two-tailed tests). t-statistics (in parentheses) are calculated based on standard errors clustered by industry-year.

Table 6. Economic Relevance

Variables	(1) CASH_ETR	(2) CASH_ETR	(3) CASH_ETR
pos_MP	-0.0187**		
	(-2.5840)		
neg_OP		-0.0200**	
		(-2.3172)	
pos_MP x neg_OP			-0.0347***
			(-3.4955)
OP	0.0722***		
	(3.7732)		
MP		-0.0008***	
		(-3.7298)	
CASH	0.1293***	0.1181***	0.1338***
	(3.4416)	(3.0993)	(3.6196)
ROA	-0.6951***	-0.6682***	-0.7004***
	(-11.1478)	(-10.5801)	(-11.2682)
LEV	0.0949***	0.0867***	0.0988***
	(4.2025)	(3.6948)	(4.4000)
Delta NOL	1.2472***	1.2533***	1.2515***
	(6.6818)	(6.7285)	(6.7150)
FI	-0.0136***	-0.0128***	-0.0141***
	(-4.4689)	(-4.2393)	(-4.6914)
PPE	0.0169	0.0159	0.0137
	(0.4907)	(0.4518)	(0.3928)
INTANG	-0.0663***	-0.0682***	-0.0653***
	(-2.6627)	(-2.7509)	(-2.6365)
EQINC	-1.7951**	-1.7296**	-1.8299**
	(-2.4064)	(-2.3260)	(-2.4228)
R&D	-0.2026*	-0.1808	-0.2131*
	(-1.6672)	(-1.5159)	(-1.7576)
SIZE	-0.0129***	-0.0067*	-0.0138***
	(-3.8432)	(-1.7485)	(-4.4228)
Delta SALE	-0.0290	-0.0374	-0.0304
	(-0.8484)	(-1.0971)	(-0.8907)
MB	-0.0004	-0.0004	-0.0004
	(-0.7372)	(-0.6614)	(-0.6208)
SG&A	0.1075***	0.1225***	0.1072***
	(3.5414)	(4.2183)	(3.4855)
Constant	0.4891***	0.4264***	0.4939***
	(10.7159)	(9.1901)	(10.4053)
Observations	4,449	4,449	4,449
Adjusted R-squared	0.1695	0.1714	0.1694
Year Dummies	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes
F	25.701	24.978	25.389

*, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively (two-tailed tests). t-statistics (in parentheses) are calculated based on standard errors clustered by industry-year.

Table 7. Robustness analysis

	Panel A:			Panel B:		
	(1) DTAX	(2) DTAX	(3) DTAX	(4) CASH_ETR	(5) GAAP_ETR	(6) DTAX
MP	0.0003**					
	(2.3120)					
OP	-0.0264***					
	(-2.6458)					
EOP		-0.0125**		0.0172	0.0282***	-0.0116*
		(-2.1159)		(1.4678)	(2.6503)	(-1.9124)
EMP		0.0002***		-0.0488*	-0.0425**	0.0264**
		(3.0703)		(-1.9506)	(-2.1620)	(2.3012)
SOP			-0.0171*			
			(-1.7945)			
SMP			0.0002*			
			(1.8610)			
CSRP without EP		0.0001				
		(0.8095)				
CSRP without SP			0.0004**			
			(2.0983)			
CSR residual				-0.0010***	-0.0005**	0.0002
				(-4.2997)	(-2.3376)	(1.4343)
Observations	2,678	2,678	2,678	4,449	4,427	2,678
Adjusted R-squared	0.4744	0.4746	0.4733	0.1709	0.0503	0.4736
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Industry Dummies	No	No	No	Yes	Yes	Yes
F	29.171	28.028	27.653	25.099	7.1989	27.909

*, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively (two-tailed tests).
t-statistics (in parentheses) are calculated based on standard errors clustered by industry-year.

References

Armstrong, C., Blouin, J., Jagolinzer, A., Larker, D., 2014. Corporate Governance, Incentives, and Tax Avoidance. Working paper, University of Pennsylvania.

Bebbington, J., Larrinaga, C., Moneva, J.M., 2008. Corporate social reporting and reputation risk management. Account. Audit. Account. J. 21(3), 337-361.

Carroll, A.B., 2015. Corporate social responsibility: The centerpiece of competing and complementary frameworks. Organ. Dyn. 44, 87-96.

Chen, S., Chen, X., Cheng, Q., Shevlin, T., 2010. Are family firms more tax aggressive than non-family firms? J. Financ. Econ. 95, 41-61.

Cheng, B., Ioannou, I., Serafeim, G., 2014. Corporate social responsibility and access to finance. Strateg. Manage. J. 35(1), 1-23.

Cho, C.H., Guidry, R.P., Hageman, A.M., Patten, D.M., 2012. Do actions speak louder than words? An empirical investigation of corporate environmental reputation. Account. Bus. Res. 37, 14-25.

Christensen, D.M., 2016. Corporate accountability reporting and high-profile misconduct. Account. Rev. 91(2), 377-399.

Davis, A.K., Guenther, D.A., Krull, L.K., Williams, B.M., 2016. Do socially responsible firms pay more taxes? Account. Rev. 91(1), 47-68.

Desai, M.A., Dharmapala, D., 2006. Corporate tax avoidance and high-powered incentives. J. Account. Econ. 79, 145-179.

Dyreng, S., Hanlon, M., Maydew, E., 2008. Long-run corporate tax avoidance. Account. Rev. 83, 61-82.

Elkington, J., 1997. Cannibals with forks: The triple bottom line of 21st century business. Capstone, Oxford.

Fombrun, C., 1996. Reputation: Realizing Value from the Corporate Image. Harvard Business School Press, Boston, MA.

Frank, M., Lynch, L., Rego, S., 2009. Tax reporting aggressiveness and its relation to aggressive financial reporting. Account. Rev. 84, 467-496.

Freeman, R.E., 1984. Strategic management: A stakeholder approach. Cambridge University Press, Boston, MA.

Gaertner, F.B., 2014. CEO after-tax compensation incentives and corporate tax avoidance. Contemp. Account. Res. 31(4), 1077-1102.

Godfrey, P., Merrill, C., Hansen, J., 2009. The relationship between corporate social responsibility and shareholder value: An empirical test of the risk management hypothesis. Strateg. Manage. J. 30, 425-445.

Gupta, S., Newberry, K., 1997. Determinants of the variability in corporate effective tax rates: evidence from longitudinal data. J. Account. Public Policy 16(1), 1-34.

Hanlon, M., Heitzman, S., 2010. A review of tax research. J. Account. Econ. 50, 127-178.

Hill, C., Jones, G., 2001. Strategic Management. Mifflin, Houghton.

Hillman, A., Keim, G.D., 2001. Shareholder value, stakeholder management, and social issues: what's the bottom line? Strateg. Manage. J. 22(125-139).

Hoi, C.K., Wu, Q., Zhang, H., 2013. Is corpoate social responsibility (CSR) associated with tax avoidance? Evidence from irresponsible CSR activities. Account. Rev. 88(6), 2025-2059.

Hummel, K., Schlick, C., 2016. The Relationship between Sustainability Performance and Sustainability Disclosure – Reconciling Voluntary Disclosure Theory and Legitimacy Theory. J. Account. Public Policy 35, 455-476.

Jones, J.J., 1991. Earnings management during import relief investigations. J. Account. Res. 29(2), 193-228.

Jung, E.J., Kim, J.S., Rhee, S.K., 2001. The measurement of corporate environmental performance and its application to the analysis of efficiency in oil industry. J. Clean. Prod. 9(6), 551-563.

Laguir, I., Stagliano, R., Elbaz, J., 2015. Does corporate social responsibility affect corporate tax aggressiveness? J. Clean. Prod. 107, 662-675.

Lanis, R., Richardson, G., 2012. Corporate social responsibility and tax aggressiveness: an empirical analysis. J. Account. Public Policy 31, 86-108.

Lanis, R., Richardson, G., 2015. Is corporate social responsibility performance associated with tax avoidance? J. Bus. Ethics 127, 439-457.

Matten, D., Moon, J., 2008. "Implicit" and "explicit" CSR: A conceptual framework for a comparative understanding of corporate social responsibility. Acad. Manage. Rev. 33(2), 404-424.

Rego, S., 2003. Tax avoidance activities of U.S. multinational corporations. Contemp. Account. Res. 20, 805-833.

Rehbein, K., Waddock, S.A., Graves, S.B., 2004. Understanding shareholder activism: which corporations are targeted? Bus. Soc. 43, 239-267.

Thomson Reuters, 2012. Asset4 ESG Data. https://www.thomsonreuters.com/content/dam/openweb/documents/pdf/tr-com-financial/fact-sheet/esg-data-fact-sheet.pdf (accessed 15.07.2017).

Thomson Reuters, 2013. Corporate responsibility ratings (TRCRR), . https://www.thomsonreuters.com/content/dam/openweb/documents/pdf/tr-com-financial/methodology/corporate-responsibility-ratings.pdf (accessed 15.07.2017).

Trumpp, C., Endrikat, J., Zopf, C., Guenther, E., 2015. Definition, conceptualization, and measurement of corporate environmental performance: A critical examination of a multidimensional construct. J. Bus. Ethics 126(2), 185-204.

Wood, D.J., 1991. Corporate social performance revisited. Acad. Manage. Rev. 16(4), 691-718.

Xie, S., Hayase, K., 2007. Corporate environmental performance evaluation: A measurement model and a new concept. Bus. Strat. Environ. 16(2), 148-168.

Ziegler, A., Busch, T., Hoffmann, V.H., 2011. Disclosed corporate responses to climate change and stock performance: An international empirical analysis. Energ. Econ. 33(6), 1283-1294.

Appendix

Table I. Variables description

Variable	Definition	Source
Measures of tax avoidance:		
CASH_ ETR	The five-year cash effective tax rate for firm i in year t, which is defined as the five-year average of cash taxes paid (TXPD) divided by the pretax book income (PI) less special items (SPI). The CASH_ETR takes a value of 1 if it exceeds the value of 1 or if the denominator is negative and cash taxes paid are positive. The CASH_ETR takes a value of 0 if cash taxes paid are negative.	Compustat
GAAP_ETR	The GAAP effective tax rate for firm i in year t, which is defined as total income taxes (TXT) divided by pretax income (PI). The GAAP_ETR takes a value of 1 if it exceeds 1 or if the pretax income is negative and total income taxes are positive. The GAAP ETR takes a value of 0 if total income taxes are negative.	Compustat
Measures of CSP:		
MP	Management performance for firm i in year t, which is measured as the residuals from the following regression: $CSRP_{i,t} = \beta_0 + \beta_1 OP_{i,t} + \epsilon_{i,t}$, where CSRP is a firm's overall corporate sustainability performance (CSP), measured as the average of its environmental (ENVSCORE) and social performance (SOCSCORE). OP is a firm's operational sustainability performance.	Asset 4
OP	Operational performance for firm i in year t, which is measured as the mean of a firm's environmental operational performance (EOP) and social operational performance (SOP).	Asset 4
EOP	Environmental operational performance for firm i in year t, which is measured as the mean of energy consumption (ENRRDP033), water withdrawal (ENRRDP054), CO_2 emissions (ENERDP023) and waste produced (ENERDP045). All items are negatively scaled, winsorized at the top and bottom 1% level by year, and standardized by three-digit GICS industry codes.	Asset 4

EMP	Environmental management performance for firm i in year t, which is measured as the residuals from the following regression: $ENVP_{i,t} = \beta_0 + \beta_1 EOP_{i,t} + \epsilon_{i,t}$, where ENVP is a firm's overall environmental performance (ENVSCORE), and EOP is a firm's environmental operational performance.	Asset 4
SOP	Social operational performance for firm i in year t, which is measured as the mean of turnover (SOEQDP034, negatively scaled), the injury rate (SOHSDP024, negatively scaled), the average yearly training hours per employee (SOTDDP018) and the percentage of female employees (SODODP017). The injury rate and the percentage of female employees are standardized by three-digit GICS industry codes, turnover and training hours at an overall level.	Asset 4
SMP	Social management performance for firm i in year t, which is measured as the residuals from the following regression: $SOCP_{i,t} = \beta_0 + \beta_1 SOP_{i,t} + \epsilon_{i,t}$, where SOCP is a firm's overall social performance (SOCSCORE), and SOP is a firm's social operational performance (SOP).	Asset 4
CSP_wo_ENVP	CSP that is not attributable to environmental performance for firm i in year t, which is measured as the residuals from the following regression: $CSP_{i,t} = \beta_0 + \beta_1 ENVP_{i,t} + \epsilon_{i,t}$, where CSP is a firm's overall CSP, i.e., (ENVSCORE+SOCSCORE)/2, and ENVP is a firm's environmental performance (ENVSCORE).	Asset 4
CSP_wo_SOCP	CSP that is not attributable to social performance for firm i in year t, which is measured as the residuals from the following regression: $CSP_{i,t} = \beta_0 + \beta_1 SOCP_{i,t} + \epsilon_{i,t}$, where CSP is a firm's overall CSP, i.e., (ENVSCORE+SOCSCORE)/2, and SOCP is a firm's social performance (SOCSCORE).	Asset 4
Control variables:		
CASH	The cash holdings of firm i in year t, which are defined as cash and marketable securities (CHE) divided by the lagged total assets (AT).	Compustat
ROA	The profitability of firm i in year t, which is defined as pretax income (PI) less extraordinary items (XI) divided by the lagged total assets (AT).	Compustat
LEV	The leverage of firm i in year t, which is defined as the total long-term debt (DLTT) divided by the lagged total assets (AT).	Compustat
Delta NOL	The change in the tax loss carryforward of firm i in year t, which is defined as the tax loss carryforward (TLCF) less the lagged tax loss carryforward (TLCF) divided by the loss carryforward (AT). If missing, TLCF is set to 0.	Compustat
FI	Foreign income for firm i in year t, which is defined foreign pretax income (PIFO) divided by	Compustat

	the total pretax income (PI). If missing, PIFO is set to 0.	
PPE	Property, plant and equipment for firm i in year t, which is defined as property, plant and equipment (PPENT) divided by the lagged total assets (AT).	Compustat
INTANG	The intangibles of firm i in year t, which are defined as intangible assets (INTAN) divided by the lagged total assets (AT).	Compustat
EQINC	Equity income in earnings of firm i in year t, which is defined as equity income in earnings (ESUB) divided by the lagged total assets (AT).	Compustat
R&D	Research and development expenses for firm i in year t, which is defined as the research and development expenses (XRD) divided by the lagged total assets (AT). If missing, R&D is set to 0.	Compustat
SIZE	The size of firm i in year t, which is defined as the natural logarithm of total assets (AT).	Compustat
Delta SALE	The change in sales for firm i in year t, which is defined as sales turnover net(SALE) less the net lagged sales turnover (SALE) divided by the lagged total assets (AT).	Compustat
MB	The market-to-book ratio for firm i in year t, which is defined as the lagged value of shares outstanding (CSHO) times the lagged closing share price at the fiscal year-end (PRCC_F) divided by the lagged book value of equity (CEQ).	Compustat
SG&A	Selling, general and administrative expenses for firm i in year t, which is defined as selling, general and administrative expenses (XSG&A) divided by the lagged total assets (TA).	Compustat
Robustness:		
DTAX	The discretionary permanent book-tax difference (Frank et al., 2009) for firm i in year t, which is defined as the residuals from the following regression estimated by industry-year (the industry classification is based on two-digit GICS codes), where all variables (including the intercept) are scaled by the lagged total assets (AT_{t-1}): $$PERMDIFF_{i,t} = \beta_0 + \beta_1 INTAN_{i,t} + \beta_2 UNCON_{i,t} + \beta_3 MI_{i,t} + \beta_4 CSTE_{i,t} + \beta_5 Delta\ NOL_{i,t} + \beta_6 PERMDIFF_{i,t-1} + \epsilon_{i,t},$$ where $$PERMDIFF_{i,t} = BI_{i,t} - \left[\left(CFTE_{i,t} + CFOR_{i,t}\right)/STR_{i,t}\right] - DTE_{i,t}/STR_{i,t}$$ BI is a firm's pretax income (PI), CFTE is a firm's current federal tax expenses (TXFED), CFOR is a firm's current foreign tax expenses (TXFO), DTE is a firm's deferred tax expenses (TXDI), STR is the statutory tax rate, INTAN is a firm's intangible assets (INTAN), UNCON is the income reported under the equity method (ESUB), MI is the minority interest income (MII), CSTE is the current state income tax expenses (TXS), and Delta NOL is the change in the tax loss carryforward (TLCF).	Compustat

	We treat missing values as Frank et al. (2009) and Hoi et al. (2013) do.	
EMP_{TR}	Environmental management performance of firm i in year t, which is measured according to Trumpp et al. (2015).	
ESG_wo_CSP	ESG performance that is not attributable to the corporate sustainability performance of firm i in year t, which is measured as the residuals from the following regression: $ESG_{i,t} = \beta_0 + \beta_1 CSP_{i,t} + \epsilon_{i,t}$, where ESG is a firm's overall ESG performance (A4IR), and CSP is a firm's corporate sustainability performance.	Asset4
ESG_wo_ENVP	ESG performance that is not attributable to environmental performance for firm i in year t, which is measured as the residuals from the following regression: $ESG_{i,t} = \beta_0 + \beta_1 ENVP_{i,t} + \epsilon_{i,t}$, where ESG is a firm's overall ESG performance (A4IR), and ENVP is a firm's environmental performance (ENVSCORE).	Asset4
ESG_wo_SOCP	ESG performance that is not attributable to the social performance of firm i in year t, which is measured as the residuals from the following regression: $ESG_{i,t} = \beta_0 + \beta_1 SOCP_{i,t} + \epsilon_{i,t}$, where ESG is a firm's overall ESG performance (A4IR), and SP is a firm's social performance (SOCSCORE).	Asset4

For regressions with the long-run cash ETR as the dependent variable, we use mean values over a five-year period for all independent variables. Values for continuous control variables are winsorized at the top and bottom 1% level.

3. Appendix

Curriculum Vitae

Personal Details

First name Last name:	Benedikt Bisig
Date of birth:	23.12.1986

Education

October 2013 - October 2017:	PhD studies at the Department of Business Administration, University of Zurich, Faculty of Business, Economics and Informatics
September 2011 - October 2013:	Master of Arts in Business Administration, University of Zurich, Faculty of Business, Economics and Informatics
September 2007 - July 2010:	Bachelor of Arts in Management, University of Fribourg, Faculty of Economics and Social Sciences

Work experience

March 2013 - October 2017:	Research Associate at the Chair of Accounting, University of Zurich
September 2011 - October 2017:	Commercial Manager at Bifert AG, Baar